

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2002, 10:12:21 ; Search time 7114.74 Seconds  
(without alignments)  
5903.176 Million cell updates/sec

Title: US-09-052-089A-7  
2007  
Sequence: 1 GTGCGGTGAGACGAATTGTG.....AAAAAAAAAAAAAAAAAAAA 2007

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hig:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_pro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hig\_hum:\*  
31: em\_hig\_inv:\*  
32: em\_hig\_other:\*  
33: em\_higo\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
-----					

1	2007	100.0	2007	9	HSU77845	U77845 Human hTRIP
2	1891.2	94.2	2002	9	BC019283	BC019283 Homo sapi
3	1891.2	94.2	2010	9	BC000310	BC000310 Homo sapi
4	1890.8	94.2	2065	6	AR072729	AR072729 Sequence
5	1890.8	94.2	2065	6	BD006990	BD006990 Modulator
6	1087	54.2	2024	10	BC017374	BC017374 Mus muscu
7	1081	53.9	1981	10	MMU77844	U77844 Mus musculu
8	521.6	26.0	3694	10	BC006929	BC006929 Mus muscu
9	260	13.0	129747	9	HS465N24	AL031432 Human DNA
10	212.8	10.6	391	11	G23637	G23637 human SFS W
11	176.4	8.8	7542	6	AX350344	AX350344 Sequence
12	176.4	8.8	163424	2	AC099330	AC099330 Homo sapi
13	130	6.5	110000	2	AL353694_2	Continuation (3 of
14	130	6.5	163424	2	AC099330	AC099330 Homo sapi
15	74	3.7	39410	5	FR165X09	AJ010317 Fugu rubr
16	65.8	3.3	7218	6	166494	166494 Sequence 14
17	60.6	3.0	203984	2	AC067854	AC067854 Homo sapi
18	59.6	3.0	318930	2	AC073495	AC073495 Mus muscu
19	59	2.9	216191	2	AC090495	AC090495 Mus muscu
20	58.8	2.9	2065	14	AF410847	AF410847 Ovine her
21	58.8	2.9	239480	2	AC079818	AC079818 Mus muscu
22	58	2.9	133661	14	U93872	U93872 Kaposi's sa
23	57.8	2.9	110000	2	LMFLCHR34_16	Continuation (17 o
24	57.6	2.9	219200	10	AL589701	AL589701 Mouse DNA
25	57	2.8	3012	14	AF360120	AF360120 Human her
26	57	2.8	192675	10	AL589767	AL589767 Mouse DNA
27	56.4	2.8	224712	2	AL606969	AL606969 Mus muscu
28	56.2	2.8	174019	2	AC027267	AC027267 Homo sapi
29	56.2	2.8	185497	2	AC073446	AC073446 Homo sapi
30	56.2	2.8	223426	2	AL606962	AL606962 Mus muscu
31	56	2.8	282611	2	AL645746	AL645746 Mus muscu
32	55.8	2.8	28559	14	AF148805	AF148805 Kaposi's
33	55.4	2.8	3127	14	AF305694	AF305694 Kaposi's
34	55.4	2.8	207683	2	AC098712	AC098712 Mus muscu
35	55	2.7	35848	2	AC011553	AC011553 Homo sapi
36	54	2.7	110000	9	AC091454_2	Continuation (3 of
37	53.8	2.7	180385	9	AC007461	AC007461 Homo sapi
38	53.8	2.7	241432	10	AL589661	AL589661 Mouse DNA
39	53.4	2.7	195294	2	AC051623	AC051623 Mus muscu
40	53.4	2.7	202010	2	AC016961	AC016961 Homo sapi
41	53.4	2.7	236685	2	AC084744	AC084744 Mus muscu
42	53.2	2.7	894	9	ATRINVA	M25314 Owl monkey
43	53.2	2.7	1635	9	ATRINVA	M25313 Owl monkey
44	53.2	2.7	227605	2	AC073564	AC073564 Mus muscu
45	53	2.6	142277	2	AC016204	AC016204 Homo sapi

## ALIGNMENTS

RESULT	LOCUS	Human hTRIP (hTRIP) mRNA, complete cds.	2007 bp	linear	PRI 25-APR-1997
1	HSU77845	U77845	U77845.1	GI:2039303	
ORGANISM Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
REFERENCE	AUTHORS	Lee,S.Y., Lee,S.Y. and Choi,Y.			
TITLE	TRAF-Interacting protein (TRIP): a novel component of the tumor necrosis factor receptor (TNFR)- and CD30-TRAF signaling complexes that inhibits TRAF2-mediated NF-kappaB activation				
JOURNAL	J. Exp. Med. 185 (7), 1275-1285 (1997)				
REFERENCE	MEDLINE	97258620			
AUTHORS	2 (bases 1 to 2007)				
TITLE	Lee,S.Y., Lee,S.Y. and Choi,Y.				
JOURNAL	Submitted (10-NOV-1996) Immunology, The Rockefeller University, 1230 York Avenue Box 295, New York, NY 10021, USA				

FEATURES  
Source Location/Qualifiers  
1..2007  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..2007  
/gene="htr1p"  
/gene="htr1p"  
103..1512  
/gene="htr1p"  
/codon\_start=1  
/product="htr1p"  
/protein\_id="AA052993.1"  
/db\_xref="GI:2039304"  
/translation="MPTRALCTICSDFFDHSRDVAIHCGHTFHLCILQSEETAPSR  
TCPOCRIOVQKRTIINKLFEDLAEBEENVLDREFLNEDNVRAQLSQDKERDSQV  
IDTLDLTLEERNATVVSLOALGKAEMLCSTLKQMKLEQOODETKOAEAGRLR  
SKMTKEOIELLOSLQPEYEMIRDMGQSAVEQLAVYCVSLKREYENLKERARAS  
GVADKLKRDDESSRSKLIQVYSELDQATLELAKSNQDKQSDMKETMSLKRLTMQOE  
TLMPPVASETDRLVLESPAPVEVNLKRPSEFDIDLNATFEDVTPPARSSOH  
GYEKLCTEKSHSPIDVPRKICGKPRKSSQLSGQSCAGEDEDLVGAFLFVRNA  
ILGOKPKRPRSSCSKDVVRFGFDLGRTKFIQPTDVMIRPLPVRPKTKVKORV  
RYKTVPSLFOAKLDTFLMS"

BASE COUNT 517 a 518 c 558 g 414 t  
ORIGIN

Query Match 100.0%; Score 2007; DB 9; Length 2007;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2007; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCGGTGGAGCCGAATTTAGACAGCGAGCGGGCCCTTACGAGCCGAGCTGT 60  
DB 1 GTGCGGTGGAGCCGAATTTAGACAGCGAGCGGGCCCTTACGAGCCGAGCTGT 60  
QY 61 AGCAGTTCTTTGGCTCCTGGGCCCCCTTGATCCAGCCATTCGCTATCCGTCGTG 120  
DB 61 AGCAGTTCTTTGGCTCCTGGGCCCCCTTGATCCAGCCATTCGCTATCCGTCGTG 120  
QY 121 TGCATATCTGCTCCGACTTCTTGATCACTCCGCGAGCTGGCCGACATCCATGCGGC 180  
DB 121 TGCATATCTGCTCCGACTTCTTGATCACTCCGCGAGCTGGCCGACATCCATGCGGC 180  
QY 181 CATACCTTCCACTTGGCAGTGCCTTAATTCAGTCTTTGAGACAGCACCAGTGGAGCTGC 240  
DB 181 CATACCTTCCACTTGGCAGTGCCTTAATTCAGTCTTTGAGACAGCACCAGTGGAGCTGC 240  
QY 241 CCACAGTGGCCGAATCCAGTGTGGCAAAAGAACCATTAATCAATGAAGCTCTTTGATCTT 300  
DB 241 CCACAGTGGCCGAATCCAGTGTGGCAAAAGAACCATTAATCAATGAAGCTCTTTGATCTT 300  
QY 301 GCCCAGAGAGGAGAGATGTCTTGATCGAGAAATTTCTTAAGATGAATGACAATGTC 360  
DB 301 GCCCAGAGAGGAGAGATGTCTTGATCGAGAAATTTCTTAAGATGAATGACAATGTC 360  
QY 361 AGAGCCAGCTTTCCAGAAAGACAGAGAGAAACGACAGCCAGGTCATCTGACACT 420  
DB 361 AGAGCCAGCTTTCCAGAAAGACAGAGAGAAACGACAGCCAGGTCATCTGACACT 420  
QY 421 CTGCGGATACGCTGGAAGAACCAATGCTACTGTGATCTCTGACAGCGGCTTGAGG 480  
DB 421 CTGCGGATACGCTGGAAGAACCAATGCTACTGTGATCTCTGACAGCGGCTTGAGG 480  
QY 481 AAGCGGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAATCTTAGAGCAGCAG 540  
DB 481 AAGCGGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAATCTTAGAGCAGCAG 540  
QY 541 GATGAGACCAACAGACAGAGAGAGGGGCGGCTCGAGAGCAAGATGAAGACCATG 600  
DB 541 GATGAGACCAACAGACAGAGAGAGGGGCGGCTCGAGAGCAAGATGAAGACCATG 600  
QY 601 GACGAGATTGAGCTTCTACTCCAGAGCCAGCTCCGAGGTGGAGAGATGATCCAGAC 660  
DB 601 GACGAGATTGAGCTTCTACTCCAGAGCCAGCTCCGAGGTGGAGAGATGATCCAGAC 660

QY 661 ATGGGTGTGGACAGTCAGCGGTGGAACAGCTGGCTGTACTGTGTCTCTCAAGAAA 720  
DB 661 ATGGGTGTGGACAGTCAGCGGTGGAACAGCTGGCTGTACTGTGTCTCTCAAGAAA 720  
QY 721 GACTAGAGATATTTAAAGAGCGCAGGAAGGCTTCAGGGAGGTGCTGACAACTGAGG 780  
DB 721 GACTAGAGATATTTAAAGAGCGCAGGAAGGCTTCAGGGAGGTGCTGACAACTGAGG 780  
QY 781 AAGGATTTGTTTCTCCAGAACCAAGTTGACAGAGTCACTCTGAATTTGATCGAGCC 840  
DB 781 AAGGATTTGTTTCTCCAGAACCAAGTTGACAGAGTCACTCTGAATTTGATCGAGCC 840  
QY 841 AAGTTAGAACTGAAGTCAGCCAGGAAGACTTCAGAGTGTCTGACAGGAATCATGAGC 900  
DB 841 AAGTTAGAACTGAAGTCAGCCAGGAAGACTTCAGAGTGTCTGACAGGAATCATGAGC 900  
QY 901 CTGAAAAAGAGCTAAACGATGCTGACAGAAACCTTAACTGCGACCAATGGCCAGTGAG 960  
DB 901 CTGAAAAAGAGCTAAACGATGCTGACAGAAACCTTAACTGCGACCAATGGCCAGTGAG 960  
QY 961 ACTGTGACCGGCTGGTTTTAGAGAGCCAGCCCTGTGGAGGTGAATCGAAGCTCCGC 1020  
DB 961 ACTGTGACCGGCTGGTTTTAGAGAGCCAGCCCTGTGGAGGTGAATCGAAGCTCCGC 1020  
QY 1021 CGGCATCTCTCCGTGATGATATTTGATCTCAATGCTACCTTGTGATGATCTCCCA 1080  
DB 1021 CGGCATCTCTCCGTGATGATATTTGATCTCAATGCTACCTTGTGATGATCTCCCA 1080  
QY 1081 GCGCGGCTCTCCAGTCTCCAGCATGTGTTACTACGAAAACTTTGCTTACAGAACTACAC 1140  
DB 1081 GCGCGGCTCTCCAGTCTCCAGCATGTGTTACTACGAAAACTTTGCTTACAGAACTACAC 1140  
QY 1141 TCCCAATTCAGATGATGCCCAAGAAATATGCAAGAGCCCAAGAAAGTCCAGCTC 1200  
DB 1141 TCCCAATTCAGATGATGCCCAAGAAATATGCAAGAGCCCAAGAAAGTCCAGCTC 1200  
QY 1201 TCACCTGGGTGGCCAGAGCTGTGACAGAGAGCCAGATGAGAACTGTTGTTGCCCT 1260  
DB 1201 TCACCTGGGTGGCCAGAGCTGTGACAGAGAGCCAGATGAGAACTGTTGTTGCCCT 1260  
QY 1261 ATTTTGTCCGGAATGCCATCTTACGAGCAAGAAACAGCCCAAGAGGCTCAGAGTGC 1320  
DB 1261 ATTTTGTCCGGAATGCCATCTTACGAGCAAGAAACAGCCCAAGAGGCTCAGAGTGC 1320  
QY 1321 TCTTGCAGAAAGATGTGGTAAAGACAGGCTTGATGGGCTGGTGGCCGCAAAATTC 1380  
DB 1321 TCTTGCAGAAAGATGTGGTAAAGACAGGCTTGATGGGCTGGTGGCCGCAAAATTC 1380  
QY 1381 ATCCAGCTACTGACACAGTCAATGATCCGCCCAATTCCTGTTAAAGCCAAAGCAAGTT 1440  
DB 1381 ATCCAGCTACTGACACAGTCAATGATCCGCCCAATTCCTGTTAAAGCCAAAGCAAGTT 1440  
QY 1441 AAGCAGAGGTGAGGTTGAAGACCGTGTCTTCTTCCAGGCCAGCTGGACACTTC 1500  
DB 1441 AAGCAGAGGTGAGGTTGAAGACCGTGTCTTCTTCCAGGCCAGCTGGACACTTC 1500  
QY 1501 CTGTGCTGAGGAACAGAGTCTGACCAATGGCCAGACACATGCTGCAACTGTAGG 1560  
DB 1501 CTGTGCTGAGGAACAGAGTCTGACCAATGGCCAGACACATGCTGCAACTGTAGG 1560  
QY 1561 TCAAGGACTGTCCAGCAGAGGTTTGTGACAGAGCCCTACTTTCGGGACACAGCTGAGGT 1620  
DB 1561 TCAAGGACTGTCCAGCAGAGGTTTGTGACAGAGCCCTACTTTCGGGACACAGCTGAGGT 1620  
QY 1621 GTAAGGGCAGACAAACAGGTGAGGTTGAGTGTACACCCAGAGACTGCTTCTGCTCCCT 1680  
DB 1621 GTAAGGGCAGACAAACAGGTGAGGTTGAGTGTACACCCAGAGACTGCTTCTGCTCCCT 1680  
QY 1681 CACCTGGCCCACTCTCTAGACTGGGAGCTGACATGACCAGCCACTGATCTGTACGA 1740  
DB 1681 CACCTGGCCCACTCTCTAGACTGGGAGCTGACATGACCAGCCACTGATCTGTACGA 1740  
QY 1741 GGTCTGCTCTGTTCGACAGGCTTGTATAGCCATGATCAGATGTGCTCAGACTCTTT 1800

Db 1741 GGCCCTGCTCTGTGGCAGGCTTGTATAGCCATGATGATGAGATGTCAGACATCTTT 1800  
Qy 1801 CTGGGCGCTGGAGACCCAGGCTCACTTGTCTCTGTGACCCAGATGCTTGAGGA 1860  
Db 1801 CTGGGCGCTGGAGACCCAGGCTCACTTGTCTCTGTGACCCAGATGCTTGAGGA 1860  
Qy 1861 TCTCAGGACGCTCAGCCAGCTTACCTGCTTGACTGCTCTAGCATAGCCTGG 1920  
Db 1861 TCTCAGGACGCTCAGCCAGCTTACCTGCTTGACTGCTCTAGCATAGCCTGG 1920  
Qy 1921 GCCAAGCAGGCTGGGGAATGAGAGATAGACATGAGATGATGAGAGATGGAATTTT 1980  
Db 1921 GCCAAGCAGGCTGGGGAATGAGAGATAGACATGAGATGATGAGAGATGGAATTTT 1980  
Qy 1981 CCCGAAAAAAAAAAAAAAAAAAAA 2007  
Db 1981 CCCGAAAAAAAAAAAAAAAAAAAA 2007

RESULT 2  
LOCUS BC019283 2002 bp mRNA linear PRI 19-DEC-2001  
DEFINITION Homo sapiens, TRAF interacting protein, clone MGC:3999  
IMAGE:2821007, mRNA, complete cds.  
ACCESSION BC019283  
VERSION BC019283.1 GI:17939476  
KEYWORDS MGC.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 2002)  
Strausberg R.  
Direct Submission  
Submitted (13-DEC-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: http://mgc.nci.nih.gov  
Contact: MGC help desk  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
info@bcgsc.bc.ca  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Lisa Pirabhu, Parvaneh Saeedi, Jacquelline  
Schlein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranda Tsai, Natasha van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
Series: IRAL Plate: 9 Row: 1 Column: 1  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 5032194.

FEATURES  
Source  
1..2002  
/organism="Homo sapiens"  
/db\_xref="LocusID:10293"  
/db\_xref="taxon:9606"  
/clone="MGC:3999 IMAGE:2821007"  
/tissue\_type="lung, small cell carcinoma"  
/clone\_1lb="NIH-MGC\_7"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
88..1497  
CDS

/codon\_start=1  
/product="TRAF interacting protein"  
/protein\_id="AAH19283.1"  
/db\_xref="GI:17939477"  
/translation="MPRALCTICSDFFDHSRYVAIHCHTFHLQCLIFOWEFTAPSR  
TCPOCRIVGKRTIINKLFPLDAGEEVNLADEFILNKLDVNRRAQLSOKDEKRSQV  
IIDLRLTEERNATVVSILQOALKAEMICSTLKKOMYLEGOODETQAOEARRLR  
SKMTBEOIELLLOSOPREVERMIPDMVGOSAVROLAVCYSLKKEVENIKEARAS  
GEVADRLKDLFPSSRSKLOTYSELDOAKLELKSOKLOSDKELMSLKKLTLLOE  
TLNLPVASETVDRVLVLESPAVEVNALKLRPSFDDIDLNATEFDVDPAPRSSOH  
GYEKLCLCEKSHSPIQDVPKICKPKRRESQSLSGOSCASEPDELGAPPIFYRNA  
ILGOKOPRPSSESSKDVMTGFDGGLGGRKFIQPDVTWIRPLPVKPKTKVQKV  
RKVTVPSLFEQKLDPELMS"  
BASE COUNT 516 a 517 c 556 g 413 t  
ORIGIN  
Query Match 94.2%; Score 1891.2; DB 9; Length 2002;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 1964; Conservative 0; Mismatches 23; Indels 9; Gaps 5;  
Qy 20 GAAGCAAGCGGAGGCGGGCG----CTTACGAGAGCGGACCTGTAGCAGTTCTTTGGC 75  
Db 1 GAAGCAAGCGGAGGCGGGCGGGCGGCGCTTACGAGAGCGGACCTGTAGCAGTTCTTTGGC 60  
Qy 76 TGGCTGGGGCCCTTGAGTGCAGGCATGAGCCATGCCGTCGTGTGCACTATCTGCTCC 135  
Db 61 TGGCTGGGGCCCTTGAGTGCAGGCATGAGCCATGCCGTCGTGTGCACTATCTGCTCC 120  
Qy 136 GACTTCTTGATCATCTCCCGACGAGTGGCGGCATCAGTGGCGGCACACCTTCCACTTG 195  
Db 121 GACTTCTTGATCATCTCCCGACGAGTGGCGGCATCAGTGGCGGCACACCTTCCACTTG 180  
Qy 196 CAGTCCCTAATTCATGTCCTTTGAGACAGACACCAAGTGGACCTGCCACAGTGGCAATC 255  
Db 181 CAGTCCCTAATTCATGTCCTTTGAGACAGACACCAAGTGGACCTGCCACAGTGGCAATC 240  
Qy 256 CAGGTTGGCAAAAGAACCAATTCATTAAGCTCTTTGATCTTCCAGAGAGAGAG 315  
Db 241 CAGGTTGGCAAAAGAACCAATTCATTAAGCTCTTTGATCTTCCAGAGAGAGAG 300  
Qy 316 AATGCTTGGATCGAATTTCTTAAGAAATGAATGAGACATGTCAGACCCAGCTTTC 375  
Db 301 AATGCTTGGATCGAATTTCTTAAGAAATGAATGAGACATGTCAGACCCAGCTTTC 360  
Qy 376 CAGAAAGACAGAGAGAGACAGACAGCAGGTGATCATGACACTCTGCGGGATACGCTG 435  
Db 361 CAGAAAGACAGAGAGAGAGACAGACAGCAGGTGATCATGACACTCTGCGGGATACGCTG 420  
Qy 436 GAAGAACGAAATGCTATGTCGATCTGACGACGCTTGGGCAAGGCCGAGATGCTG 495  
Db 421 GAAGAACGAAATGCTATGTCGATCTGACGACGCTTGGGCAAGGCCGAGATGCTG 480  
Qy 496 TGGTCCACTGAAAGAGAGATGAGTACTTATGACACAGCAGAGATGAGACAAACA 555  
Db 481 TGGTCCACTGAAAGAGAGATGAGTACTTATGACACAGCAGAGATGAGACAAACA 540  
Qy 556 GCACAAGAGAGAGGCGGGCGGCTCAGAGACCAAGATGAAGACATGAGAGATTTAGCTT 615  
Db 541 GCACAAGAGAGAGGCGGGCGGCTCAGAGACCAAGATGAAGACATGAGAGATTTAGCTT 600  
Qy 616 CTACTCCAGACAGCAGCTCCCTGAGTGGAGAGATATCCGAGACATGGGTGTGGACAG 675  
Db 601 CTACTCCAGACAGCAGCCTCGTGGTGGAGAGATATCCGAGACATGGGTGTGGACAG 660  
Qy 676 TCAGCGGTGAACAGCTGCTGTACTGTGTCTCTCAAGAAAGATACGAGATCTA 735  
Db 661 TCAGCGGTGAACAGCTGCTGTACTGTGTCTCTCAAGAAAGATACGAGATCTA 720  
Qy 736 AAAGAGCAGCGAAGGCTCAGGGGAGGTGGCTGACAACTGAGAGAGATTTGTTTCC 795  
Db 721 AAAGAGCAGCGAAGGCTCAGGGGAGGTGGCTGACAACTGAGAGAGATTTGTTTCC 780

QY 796 TCAGAGCAAGATTGACAGAGTCTACTGTGAATTGGATGAGGCCAAGTAGAATGAG 855  
 |||||  
 Db 781 TCAGAGCAAGATTGACAGAGTCTACTGTGAATTGGATGAGGCCAAGTAGAATGAG 840  
 QY 856 TCAGAGCAAGATTGACAGAGTCTACTGTGAATTGGATGAGGCCAAGTAGAATGAG 915  
 |||||  
 Db 841 TCAGAGCAAGATTGACAGAGTCTACTGTGAATTGGATGAGGCCAAGTAGAATGAG 900  
 QY 916 ACAGTGTGAGAGCAAGATTGACAGAGTCTACTGTGAATTGGATGAGGCCAAGTAGA 975  
 |||||  
 Db 901 ACAGTGTGAGAGCAAGATTGACAGAGTCTACTGTGAATTGGATGAGGCCAAGTAGA 960  
 QY 976 GTTTTAGAGAGCCAGCCCTGTGAGAGTGAATCTGAAGCTCCGCCGCCATCCTTCGT 1035  
 |||||  
 Db 961 GTTTTAGAGAGCCAGCCCTGTGAGAGTGAATCTGAAGCTCCGCCGCCATCCTTCGT 1020  
 QY 1036 GATGATATTTATCTCATCTGCTACTCTTGTGATGTGATATCTCCCGAGCCGCCCTCAGC 1095  
 |||||  
 Db 1021 GATGATATTTATCTCATCTGCTACTCTTGTGATGTGATATCTCCCGAGCCGCCCTCAGC 1080  
 QY 1096 TCCAGAGCATGTTACTACGAAAACTTGGCTAGAGAGTACACACTCCCAATTCAAGAT 1155  
 |||||  
 Db 1081 TCCAGAGCATGTTACTACGAAAACTTGGCTAGAGAGTACACACTCCCAATTCAAGAT 1140  
 QY 1156 GTCCCAAGAGATATGCAAAAGGCCCCAGAGAGAGTCCAGCTCTCATCTGGGTGGCCAG 1215  
 |||||  
 Db 1141 GTCCCAAGAGATATGCAAAAGGCCCCAGAGAGAGTCCAGCTCTCATCTGGGTGGCCAG 1200  
 QY 1216 AGCTGTCCAGAGAGAGCCAGATGAGAGAACTGGTGGTCCCTTCCCTATTTTGTCCGGAAT 1275  
 |||||  
 Db 1201 AGCTGTCCAGAGAGAGCCAGATGAGAGAACTGGTGGTCCCTTCCCTATTTTGTCCGGAAT 1260  
 QY 1276 GCCATCTTAGAGCCAGAAACAGCCCAAAAGGCCAGGTCAGAGTCCCTTTCAGCAAAAGAT 1335  
 |||||  
 Db 1261 GCCATCTTAGAGCCAGAAACAGCCCAAAAGGCCAGGTCAGAGTCCCTTTCAGCAAAAGAT 1320  
 QY 1336 GTGGTAAGAGCAGAGCTTCGATGGGCTCGGTGGCCGACAAATTCATCCAGCCTACTGAC 1395  
 |||||  
 Db 1321 GTGGTAAGAGCAGAGCTTCGATGGGCTCGGTGGCCGACAAATTCATCCAGCCTACTGAC 1380  
 QY 1396 ACAGTATGATTCGGCCATTCGCTGTTAAGCCCAAGAGTGAAGCAGAGGAGTGAGG 1455  
 |||||  
 Db 1381 ACAGTATGATTCGGCCATTCGCTGTTAAGCCCAAGAGTGAAGCAGAGGAGTGAGG 1440  
 QY 1456 GTGAAGACCTGCTCTCTCTTCACAGGCCAAGCTGGACACTTCCCTGGTGGTGAGAA 1515  
 |||||  
 Db 1441 GTGAAGACCTGCTCTCTCTTCACAGGCCAAGCTGGACACTTCCCTGGTGGTGAGAA 1500  
 QY 1516 CAGTGAAGTGCACCAATGGCCAGACACATGCTGCAACTTGTAGGTCAAGAGTGTCCAG 1575  
 |||||  
 Db 1501 CAGTGAAGTGCACCAATGGCCAGACACATGCTGCAACTTGTAGGTCAAGAGTGTCCAG 1560  
 QY 1576 GCAGGG--TTTGTGACAGAGCCCTACTTTCGGGACAGCCTGAGGTGAAGGGCAGACA 1633  
 |||||  
 Db 1561 GCAGGGGTTTGTGACAGAGCCCTACTTTCGGGACAGCCTGAGGTGAAGGGCAGACA 1620  
 QY 1634 AACAGGTGAGAGGTGAGTGTGACACCCAGAGAGTCTCTTCTCCCTCACCCTGCCCCAC 1693  
 |||||  
 Db 1621 AACAGGTGAGAGGTGAGTGTGACACCCAGAGAGTCTCTTCTCCCTCACCCTGCCCCAC 1680  
 QY 1694 TCTTACAGATGAGAGCTGACATGACAGCCAGCCACTGATCTGTGACAGAGTGTCTGCT 1752  
 |||||  
 Db 1681 TCTTACAGATGAGAGCTGACATGACAGCCAGCCACTGATCTGTGACAGAGTGTCTGCT 1740  
 QY 1753 TTGCGAGGGCTGTTTATAGCCATGATCAGATGAGTGTGACAGTCTTTCGAGGCTGAGAG 1812  
 |||||  
 Db 1741 TTGCGAGGGCTGTTTATAGCCATGATCAGATGAGTGTGACAGTCTTTCGAGGCTGAGAG 1800  
 QY 1813 ACCAGGCTGACTTGTGACTGTCTGTGGACAGAGAGTCTTTCGAGGCTGAGAGTCTGAGCAGCC 1872  
 |||||  
 Db 1801 ACCAGGCTGACTTGTGACTGTCTGTGGACAGAGAGTCTTTCGAGGCTGAGAGTCTGAGCAGCC 1860  
 QY 1873 TCAGCCCAAGCTTCTACTGCTTGTGACTTGTCTTA--GCATAGCTGTGGGCCAAGAGAGG 1931

Db 1861 TCAGCCCAAGCTTCTACTGCTCTTGTAGGATGAGCCCTGGGCCAAGCAGAGG 1920  
 QY 1932 TGGGGAATGAGAGATGACATGAGATGTATGAGAGATGAGGAAGATTTTCCGAAAAA 1991  
 |||||  
 Db 1921 TGGGTAATGAGAGATG--CATGGATGTATGAGAGATGAGGAATTTTCATGTAAATA 1979  
 QY 1992 AAAAAAAAAAAAAA 2007  
 |||||  
 Db 1980 AAAAAAAAAAAAAA 1995

RESULT 3  
 BC000310 2010 bp mRNA linear PRI 12-JUL-2001  
 LOCUS Homo sapiens, TRAF interacting protein, clone MGC:8424  
 DEFINITION IMAGE:2821007, mRNA, complete cds.  
 ACCESSION BC000310  
 VERSION BC000310.1 GI:12653088  
 KEYWORDS MGC.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 2010)  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-NOV-2000) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 COMMENT Contact: MGC help desk

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: DCTD/DTF  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland.  
 Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgc@nrl.nih.gov](mailto:nisc_mgc@nrl.nih.gov)  
 Shevchenko, Y., Weltherby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantrilpop, S., Thomas, P.J.,  
 Tlionson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.nih.gov>  
 Series: IRL Plate: 1 Row: e Column: 7  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 5032194.

#### FEATURES

source

#### CDS

1..2010  
 /organism="Homo sapiens"  
 /db\_xref="LocusID:10293"  
 /db\_xref="taxon:9606"  
 /clone="MGC:8424 IMAGE:2821007"  
 /tissue\_type="lung, small cell carcinoma"  
 /clone\_id="NIH MGC 7"  
 /lab\_host="DH10B-R"  
 /note="vector: pOTB7"  
 97..1506  
 /codon\_start=1  
 /product="TRAF interacting protein"  
 /protein\_id="AAH00310.1"  
 /db\_xref="GI:12653088"  
 /translation="MPRALCTTSDPFDHSDVAAIHGTHPHLOCLIQWETAPSR  
 TPCQRIQVGRRTIINKLFEDLAEENEVLDLAEFLKNELDVNRALDSQDEKRDQV  
 IIDLRLDLERNAIVVSLDQALGKAEMLCSTLKKOMKYLEBOODDET KOADEFARLR

SKMKTMTIOELLLOSORPEVEEMTRDMGVGSAVEOLAVYCSKREYENIKREKRS  
 GEVMDKRLKDESSRSKLTQYVSELDQALIKSAOKDLOSADKEIMSKIKLMO  
 TYNPLVASEETDRVLVESPAPEVNNILKRPSPRDDIDNATADVTVPAPRSPSSO  
 GYKYLICKESHSDIOWPKIKCKGPKRESLSLGGSCSGEPELTVGAEPYVRNNA  
 ILGOKORPRRSPSSCSKDVVRTOFGDGLGCRTRKLTQTPDYMNRPLPVKRTKVKQSV  
 RKTVPSPFOAKLDTPLMS"

BASE COUNT	517 a	519 c	561 g	413 t
ORIGIN				

Query Match	94.2%;	Score 1891.2;	DB 9;	Length 2010;
Best Local Similarity	98.4%;	Pred. No. 0;		
Matches 1964; Conservative	0;	Mismatches 23;	Indels 9;	Gaps 5;

OY	20	GAGCAACCGGAGGCGGGGGC----	CTCTACGAGCCGGACGCTGACACTTCTTTGGC	75
Db	10	GAGCAACCGGAGGCGGGGGCCGCTCT	TGAAAGCCGGACCTGTACACTTCTTTGGC	69
OY	76	TGCGTGGGCCCTTGAGTCCAGCCCAT	TGCTATCCGTCCTGTGACACTATCTGCTC	135
Db	70	TGCGTGGGCCCTTGAGTCCAGCCCAT	TGCTATCCGTCCTGTGACACTATCTGCTC	129
OY	136	GACTTTCCTGCATCACTCCCGGAGAG	CTGCGCCATTCACACTTCCTCACTTG	195
Db	130	GACTTTCCTGCATCACTCCCGGAGAG	CTGCGCCATTCACACTTCCTCACTTG	189
OY	196	CAGTGCCTAATTCAGTCCCTTTGAGAC	AGCACCAAGTGGACCTGCCACAGTGGCCGATC	255
Db	190	CAGTGCCTAATTCAGTCTGTTTGAGAC	AGCACCAAGTGGACCTGCCACAGTGGCCGATC	249
OY	256	CAGTTCGGCAAAAGACACATTATCAT	TAAGCTCTTCTTTGATCTTGGCCAGAGAGAG	315
Db	250	CAGTTCGGCAAAAGACACATTATCAT	TAAGCTCTTCTTTGATCTTGGCCAGAGAGAG	309
OY	316	AATGCTCTGGATTCGAGATTTCTTAA	GAATGAACGTGCACATGTCACAGCCGACCTTCC	375
Db	310	AATGCTCTGGATTCGAGATTTCTTAA	GAATGAATGAACTGGACATGTCACAGCCGACCTTCC	369
OY	376	CAGAAAGCAAGAGAGAAACGAGACAG	CAGGCTCATCTGCACACTCTGCGCGATACGCTG	435
Db	370	CAGAAAGCAAGAGAGAAACGAGACAG	CAGGCTCATCTGCACACTCTGCGCGATACGCTG	429
OY	436	GAAACACCCATGCTACTGTGTGTCTCT	GACAGAGCCCTTTGGGCAAGCCGAGATGCTG	495
Db	430	GAAAGACCCATGTCTACTGTGTGTCTCT	GACAGAGCCCTTTGGGCAAGCCGAGATGCTG	489
OY	496	TGCTCCACACTGAAAAAGACAGATCAA	TACTTAGACAGCAGATGTAGACCAACAA	555
Db	490	TGCTCCACACTGAAAAAGACAGATCAA	TACTTAGACAGCAGATGTAGACCAACAA	549
OY	556	GCACAGAGAGAGCGCGCGCGCTCAG	AGCAAGATGAAGCCATGAGAGCAGATTTGAGCTT	615
Db	550	GCACAGAGAGAGAGCGCGCGCGCTCAG	AGCAAGATGAAGCCATGAGAGCAGATTTGAGCTT	609
OY	616	CTACTCCGAGACCCAGGCTCCCTGAG	AGTGGAGATGTATCCGACACATGGGTTGGGACAG	675
Db	610	CTACTCCGAGACCCAGGCTCCCTGAG	AGTGGAGATGTATCCGACACATGGGTTGGGACAG	669
OY	676	TCACGCGGTGGAACAGCTGGCTGTGT	GTACTGTGTCTCTCAAGAAAGAGTACGAGAACTTA	735
Db	670	TCACGCGGTGGAACAGCTGGCTGTGT	GTACTGTGTCTCTCAAGAAAGAGTACGAGAACTTA	729
OY	736	AAAAGAGCAGCGAAAGGCTCAGGGG	AGAGTGGCTGACAAAGTGAAGAAAGATTTGTGTTTCC	795
Db	730	AAAAGAGCAGCGAAAGGCTCAGGGG	AGAGTGGCTGACAAAGTGAAGAAAGATTTGTGTTTCC	789
OY	796	TCCAGAAAGCAGTGTGCAGACAGTGT	ACTTCTGAAATGTGATCAGGCGCAAGTATAGAACTGAG	855
Db	790	TCCAGAAAGCAGTGTGCAGACAGTGT	ACTTCTGAAATGTGATCAGGCGCAAGTATAGAACTGAG	849
OY	856	TCAGCCCAAGAGAACTTACAGAGTGT	CTGACAAAGAAATCATGAGCCTTGAAAAAGAAAGCTA	915
Db	850	TCAGCCCAAGAGAACTTACAGAGTGT	CTGACAAAGAAATCATGAGCCTTGAAAAAGAAAGCTA	909

QY	916	ACGATGCTGCAGGAAACACCTTGAAACCTGACCACACAGTGGCCAGTGAAGACTGTGACGCCCTG	975
Db	910	ACGATGCTGCAGGAAACCTTGAAACCTGACCACAGTGGCCAGTGAAGACTGTGACGCCCTG	969
QY	976	GTTTATAGAGACCCACGACCCCTGTGGAGGTGAATCTGAACCTCGCGCGGACATCTTCCTG	1035
Db	970	GTTTATAGAGACCCACGACCCCTGTGGAGGTGAATCTGAACCTCGCGCGGACATCTTCCTG	1029
QY	1036	GATGATATTGATCTCAATGCTACCTTTGATGTGATACCTCCCCACAGCCCGGCCCTCCAGC	1095
Db	1030	GATGATATTGATCTCAATGCTACCTTTGATGTGATACCTCCCCACAGCCCGGCCCTCCAGC	1089
QY	1096	TCCACAGATGGTTTACTGAGAAAACTTTGGCTGTAGAACTCACTCCCAATTACAGAT	1155
Db	1090	TCCACAGATGGTTTACTGAGAAAACTTTGGCTGTAGAACTCACTCCCAATTACAGAT	1149
QY	1156	GTCGCCAAGAAGATATGCAAAAGGCCCCAGAGAAAGATGCCACCTGCACTGGATGGCCAG	1215
Db	1150	GTCGCCAAGAAGATATGCAAAAGGCCCCAGAGAAAGATGCCACCTGCACTGGATGGCCAG	1209
QY	1216	AGCTGTGAGGAGAACGCCAGATGAGAACTGTGTGCTTCCCTATTTTGTCCGGAT	1275
Db	1210	AGCTGTGAGGAGAACGCCAGATGAGAACTGTGTGCTTCCCTATTTTGTCCGGAT	1269
QY	1276	GCCATCTCAGGCCGAAACAGCCCAAAAGGCCAGGTCAAGAGCTCTGCGCAAAAT	1335
Db	1270	GCCATCTCAGGCCGAAACAGCCCAAAAGGCCAGGTCAAGAGCTCTGCGCAAAAT	1329
QY	1336	GTGTTAAGACACGCTTCGATGGGGCTGGTGCGGAGACAAATTCATCCAGCTACTGAC	1395
Db	1330	GTGTTAAGACACAGGCTTCGATGGGGCTGGTGCGGAGACAAATTCATCCAGCTACTGAC	1389
QY	1396	ACAGCATGATCCGCCCATTTGSCGTAAAGCCCAAGCAAGATTAAGAGAGGGGTGAG	1455
Db	1390	ACAGTCATGATCCGCCCATTTGSCGTAAAGCCCAAGCAAGATTAAGAGAGGGGTGAG	1449
QY	1456	GTGAAGACCGTCCCTTCCTCTTCACAGGCCAAGCTGGACCTTCCTGTGTCGTGAA	1515
Db	1450	GTGAAGACAGTCCCTTCCTCTTCACAGGCCAAGCTGGACCTTCCTGTGTCGTGAA	1509
QY	1516	CAGTAGAGCTGACCAATGGGCCACACACATGCTCGCACTTTGATAGTCAAGACATCTCAG	1575
Db	1510	CAGTAGAGCTGACCAATGGGCCACACACATGCTCGCACTTTGATAGTCAAGACATCTCAG	1569
QY	1576	GCAAGG - TTGTGTGACAGAGCCCTTACTTTCGGGACACACCTGAGAGTGAAGGACACA	1633
Db	1570	GCAAGGGTTTTGTGTGACAGAGCCCCACTTTCGGGACACACCTGAGAGTGAAGGACACA	1629
QY	1634	AACAGGTGAGGGGTGATGTGACACCCAGAGACTGCTTCTTCGCCCTCACCCCTGCCAC	1693
Db	1630	AACAGGTGAGGATGATGTGTGGACACCCAGAGACTGCTTCTTCGCCCTCACCCCTGCCAC	1689
QY	1694	TCCATACGATGGGAGGTGACATGACACAGCCCACTGATCTCTGTACAGAGTCTCTGT	1752
Db	1690	TCCATACGATGGGAGGTGACATGACACAGCCCACTGATCTCTGTACAGAGTCTCTGT	1749
QY	1753	TTTGACAGCTCTTGTTTATAGCCATGATCAGATGTGTGAGACTCTTTCTGGGCTGTGAG	1812
Db	1750	TTTGACAGCTCTCTGTTTATAGCCATGATCAGATGTGTGAGACTCTTTCTGGGCTGTGAG	1809
QY	1813	ACCAAGGTCACTTGTATACGTCTCTGTGTGACCAAGAGTCTTGAGGACATCTCAGGACCC	1872
Db	1810	ACCAAGGTCACTTGTGTACGTCTCTGTGTGACCAAGAGTCTTGAGGACATCTCAGGACCC	1869
QY	1873	TCACGCCCAAGCTTCTACCTCCCTTGTACCTGCTTCA - GCATAGGCTGGGCCAACAGAG	1931
Db	1870	TCACGCCCAAGCTTCTACCTCCCTTGTACCTGCTTCAAGGCATAGCTGGGCCAACAGAG	1929
QY	1932	TGGGGAATGAGAGATAGACATGGGATGTATGAGAGGATGGAAGATTTTCCCGAAAAAA	1991
Db	1930	TGGGTAATGAGAGATAG - CATGGATGTATGAGAGAGATGGAAGATTTTCAATGMAATA	1988

QY 1992 AAAAAAAAAAAAAA 2007  
Db 1989 AAAAAAAAAAAAAA 2004

RESULT 4  
AR072729 2065 bp DNA linear PAT 28-AUG-2000  
LOCUS AR072729  
DEFINITION Sequence 1 from patent US 5948643.  
ACCESSION AR072729  
VERSION AR072729.1 GI:9999493  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2065)  
AUTHORS Rubinfield,B., Polakis,P.G., Lingenfelter,C. and Vuong,T.T.  
TITLE Modulators of BRCA1 activity  
JOURNAL Patent: US 5948643-A 1 07-SEP-1999;  
FEATURES Location/Qualifiers  
source 1.2065  
BASE COUNT 561 a 526 c 561 g 417 t  
ORIGIN

Query Match 94.2% Score 1890.8; DB 6; Length 2065;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1946; Conservative 0; Mismatches 17; Indels 5; Gaps 4;

QY 44 TACGAAGCCGAGACCTTACAGTTTCTTGGCTGCTGGGCCCTTGAGTCCAGCCATCA 103  
Db 44 TACGAAGCCGAGACCTTACAGTTTCTTGGCTGCTGGGCCCTTGAGTCCAGCCATCA 103

QY 104 TGGCTATCCGCTGCTGTGTGACATATCTGCTCCGACTTCTTCCATCACTCCCGGACGTGG 163  
Db 104 TGGCTATCCGCTGCTGTGTGACATATCTGCTCCGACTTCTTCCATCACTCCCGGACGTGG 163

QY 164 CCGCCATCCACTGCGGCGCCACACCTTCCACTGCTGAGTGCCTTAATTCAGTCTCTTGAAGACAG 223  
Db 164 CCGCCATCCACTGCGGCGCCACACCTTCCACTGCTGAGTGCCTTAATTCAGTCTCTTGAAGACAG 223

QY 224 CACCAAGTGGAGACCTGCGCCACAGTCCGAATCCAGTGGCAAAAGAACCATTTATCAATA 283  
Db 224 CACCAAGTGGAGACCTGCGCCACAGTCCGAATCCAGTGGCAAAAGAACCATTTATCAATA 283

QY 284 AGCTCTCTTTGATCTTGGCCAGAGAGAGAAATCTTTGGATCCAGAAATCTTTAAAGA 343  
Db 284 AGCTCTCTTTGATCTTGGCCAGAGAGAGAAATCTTTGGATCCAGAAATCTTTAAAGA 343

QY 344 ATGAACGTGACAAATGTCAGAGCCACGCTTCCAGAAAGCAAGAGAGAAACGAGACAGCC 403  
Db 344 ATGAACGTGACAAATGTCAGAGCCACGCTTCCAGAAAGCAAGAGAGAAACGAGACAGCC 403

QY 404 AGGTCAATCATGACACTCTGCGGGATACGCTGGAAAGCAATGCTTACTGTGTATCTC 463  
Db 404 AGGTCAATCATGACACTCTGCGGGATACGCTGGAAAGCAATGCTTACTGTGTATCTC 463

QY 464 TGCAGACAGCCCTTGGGCAAGGCCGAGATGCTGTCTCCACTGTAAGAAAGCAGATCAACT 523  
Db 464 TGCAGACAGCCCTTGGGCAAGGCCGAGATGCTGTCTCCACTGTAAGAAAGCAGATCAACT 523

QY 524 ACTTAGAGCAGCAGCAGATGAGACCAACAAAGCAAGAGAGAGGCGGCGCTCAGGA 583  
Db 524 ACTTAGAGCAGCAGCAGATGAGACCAACAAAGCAAGAGAGAGGCGGCGCTCAGGA 583

QY 584 GCAAGATGAAGACCATGAGCAGATTTGAGCTTCTACTCCAGAGCCAGCGCCCTGAGGTGG 643  
Db 584 GCAAGATGAAGACCATGAGCAGATTTGAGCTTCTACTCCAGAGCCAGCGCCCTGAGGTGG 643

QY 644 AGGAGATGATCCGAGACATGGGTGTGGACATGACAGCGGTGGAACAGCTGCTGTACT 703  
Db 644 AGGAGATGATCCGAGACATGGGTGTGGACATGACAGCGGTGGAACAGCTGCTGTACT 703

QY 704 GTGTGTCTCTCAGAAAGAGTACGAAATCTAAAGAGCAGGAGGCTCAGGGAGG 763  
Db 704 GTGTGTCTCTCAGAAAGAGTACGAAATCTAAAGAGCAGGAGGCTCAGGGAGG 763

QY 764 TGGCTGACAAGCTGAGAGAGATTTGTTTCTCCAGAAAGCAAGTTGACAGAGTCTACT 823  
Db 764 TGGCTGACAAGCTGAGAGAGATTTGTTTCTCCAGAAAGCAAGTTGACAGAGTCTACT 823

QY 824 CTGAATTGGATTCAGGCCCAAGTTGAAACTGAACTGACGCCAGAGAAAGCACTTACAGATGCTG 883  
Db 824 CTGAATTGGATTCAGGCCCAAGTTGAAACTGAACTGACGCCAGAGAAAGCACTTACAGATGCTG 883

QY 884 ACAAGAAATCATGAGACCTGAAAGAAAGAGTAACTGCTGCGAGAAACCTTGAACCTCG 943  
Db 884 ACAAGAAATCATGAGACCTGAAAGAAAGAGTAACTGCTGCGAGAAACCTTGAACCTCG 943

QY 944 CACCAAGTGGCAGTGAAGTGTGACCGCCCTGGTTTGAAGAGCCAGCCCTGTGGAG 1003  
Db 944 CACCAAGTGGCAGTGAAGTGTGACCGCCCTGGTTTGAAGAGCCAGCCCTGTGGAG 1003

QY 1004 TGAATCTGAAGCTCCGCGGCGCATTCCTCGTGATGATATTGATCTCAATGCTTACTTGG 1063  
Db 1004 TGAATCTGAAGCTCCGCGGCGCATTCCTCGTGATGATATTGATCTCAATGCTTACTTGG 1063

QY 1064 ATGTGATATCTCCCGGAGCCGCGGCGCTCCAGCTCCAGCATGTTACTACGAAAACTTT 1123  
Db 1064 ATGTGATATCTCCCGGAGCCGCGGCGCTCCAGCTCCAGCATGTTACTACGAAAACTTT 1123

QY 1124 GCCTAGAGAAGTCAACACTCCCAATTCAGAGATGTCCCAAGAAAGATATCAAAAGGCCCA 1183  
Db 1124 GCCTAGAGAAGTCAACACTCCCAATTCAGAGATGTCCCAAGAAAGATATCAAAAGGCCCA 1183

QY 1184 GGAAGAGTCCCAAGCTCTCACTGCTGCGCAGAGCTGTCCAGAGAGCCAGATGAGAAAC 1243  
Db 1184 GGAAGAGTCCCAAGCTCTCACTGCTGCGCAGAGCTGTCCAGAGAGCCAGATGAGAAAC 1243

QY 1244 TGGTTGGTGCCTTCCCTATTTTGTGCGGAATGCGATCTAGGCCCGAAAGCCGCAAAA 1303  
Db 1244 TGGTTGGTGCCTTCCCTATTTTGTGCGGAATGCGATCTAGGCCCGAAAGCCGCAAAA 1303

QY 1304 GGCAGAGTCAAGTCTCTTGCAGCAAGATGTGTGAAGAGCAGGCTTGTGATGGGCTCG 1363  
Db 1304 GGCAGAGTCAAGTCTCTTGCAGCAAGATGTGTGAAGAGCAGGCTTGTGATGGGCTCG 1363

QY 1364 GTGGCCGGAACAAATTCATCCAGCTTACTGACACAGTCAATGATCCGCCATTTGCTGTTA 1423  
Db 1364 GTGGCCGGAACAAATTCATCCAGCTTACTGACACAGTCAATGATCCGCCATTTGCTGTTA 1423

QY 1424 AGCCCAAGACCAAGGTTAAGCAGAGAGGTGAGGTTGAAGACCGTGCCTTCTCTCCAGG 1483  
Db 1424 AGCCCAAGACCAAGGTTAAGCAGAGAGGTGAGGTTGAAGACCGTGCCTTCTCTCCAGG 1483

QY 1484 CCAAGCTGACACCTTCTGTGTGTGTGAGAACAGTGAAGTGTGACCAATGAGCCAGACACA 1543  
Db 1484 CCAAGCTGACACCTTCTGTGTGTGTGAGAACAGTGAAGTGTGACCAATGAGCCAGACACA 1543

QY 1544 TGCCTGCAACTTGTAGGTAAAGAGTGTCCAGGACAGG--TTTGTGAGACAGCCCTACT 1601  
Db 1544 TGCCTGCAACTTGTAGGTAAAGAGTGTCCAGGACAGG--TTTGTGAGACAGCCCTACT 1601

QY 1602 TTTCGGAGCAGCCTGAGGTGAAGGCAAGCAACAGGAGGAGGTGAGTGAACCCCG 1661  
Db 1602 TTTCGGAGCAGCCTGAGGTGAAGGCAAGCAACAGGAGGAGGTGAGTGAACCCCG 1661

QY 1662 AGACTGCTCTTCTGACCCCTACCTGCTTACAGTGGAGAGCTGACATGACAG 1721  
Db 1662 AGACTGCTCTTCTGACCCCTACCTGCTTACAGTGGAGAGCTGACATGACAG 1721

QY 1722 CCCACTGATCTGTGACAGAGGTCTGCT--CTGTGCGAGGCTCTTGTATTAGCCATGAT 1780  
Db 1722 CCCACTGATCTGTGACAGAGGTCTGCT--CTGTGCGAGGCTCTTGTATTAGCCATGAT 1780

QY 1780 CCCACTGATCTGTGACAGAGGTCTGCT--CTGTGCGAGGCTCTTGTATTAGCCATGAT 1783  
Db 1780 CCCACTGATCTGTGACAGAGGTCTGCT--CTGTGCGAGGCTCTTGTATTAGCCATGAT 1783



QY	1781	CAATGTGGTACAGCTCTTTTCGGGCGCTGGACACACAGGTACCTTTGACGTCTCTGT	1840
DB	1784	CAGATGTGGTACAGCTCTTTTCGGGCGCTGGACACACAGGTACCTTTGACGTCTCTGT	1843
QY	1841	GGACACAGAGTGGTATAGGACATCTCAGGCGACCTTCAGGCCCAACCTTCTACCTGCTTGGAC	1900
DB	1844	GGACACAGAGTGGTATAGGACATCTCAGGCGACCTTCAGGCCCAACCTTCTACCTGCTTGGAC	1903
QY	1901	TTGCTTTCTA-GCATAGCCTGGGCGCAAGCAGGCTGGGGAATGAGAGATAGACATGGAGATGT	1959
DB	1904	TTGCTTTCTAAGCATATAGCCTGGGCGCAAGCAGGCTGGGGAATGAGAGATAG-CATGGAGATGT	1962
QY	1960	ATGGAGAGAGATGGAGAGATTTTCCGCAAAAAAAAAAAAAAAAAAAAAA 2007	
DB	1963	ATGGAGAGAGATGGAGAGATTTTCATGTATAAATAAAAAAAAAAAAAA 2010	
RESULT 5			
LOCUS	BD006990	2065 bp	DNA linear PAT 31-JAN-2007
DEFINITION	Modulators of BRCA1 activity.		
ACCESSION	BD006990.1	GI:18635361	
VERSION	JP 2001502893-A/1.		
KEYWORDS	unidentified.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 2065)		
AUTHORS	Luhlfeld, B., Porakisu, P., Rigenferuta, C., and Buon, T.T.		
TITLE	Modulators of BRCA1 activity		
JOURNAL	Patent: JP 2001502893-A 1 06-MAR-2001;		
COMMENT	ONIX PHARMACEUTICALS INC		
	OS Unidentified		
	PN JP 2001502893-A/1		
	PD 06-MAR-2001		
	PF 06-AUG-1997 JP 1998512659		
	PR 04-SEP-1996 US 60/025601		
	PI BONY LUBINFELD, PAUL PORAKISU, CAROL RIGENFERUTA, TERIRIN T BUON		
	PC C12N15/09, A61K38/00, A61P43/00, C07K14/435, C07K16/18, C12N1/15,		
	PC C12N1/19,		
	PC C12N1/21, C12N5/10, C12P21/02, C12P21/08, C12N15/00, C12N5/00, PC		
A61K37/02			
FEATURES			
source			
CC Strandedness: Double;			
CC Topology: Linear;			
FH Key			
FT CDS			
Location/Qualifiers			
1..2065			
/organism="unidentified"			
/db_xref="taxon:32644"			
BASE COUNT	561 a	526 c	561 g 417 t
ORIGIN			
Query Match			
Best Local Similarity 94.2%; Score 1890.8; DB 6; Length 2065;			
Matches 1946; Conservative 0; Mismatches 17; Indels 5; Gaps 4			
QY	44	TACGAAGCCGGACCTGTAGCACTTCTTTGGCTGGGCGCCCTTGAATCCAGCCATCA 103	
DB	44	TACGAAGCCGGACCTGTAGCACTTCTTTGGCTGGGCGCCCTTGAATCCAGCCATCA 103	
QY	104	TGCGATCCGCTGCTGTGACATATCTGCTGCAGCTTTCATCATCATCCCGCAGAGTG 163	
DB	104	TGCGATCCGCTGCTGTGACATATCTGCTGCAGCTTTCATCATCATCCCGCAGAGTG 163	
QY	164	CCGCATCCACATGCGGCGCACACCTTCCATTCAGATGCCCTAATTCAGTCTTGAGACAG 223	
DB	164	CCGCATCCACATGCGGCGCACACCTTCCATTCAGATGCCCTAATTCAGTCTTGAGACAG 223	
QY	224	CACCAAGTCGGACCTGCGCCACAGATGCCGAATCCAGGTGGCAAAAGACATTATCAATA 283	
DB	224	CACCAAGTCGGACCTGCGCCACAGATGCCGAATCCAGGTGGCAAAAGACATTATCAATA 283	

QY	284	AGCTCTTCTTGTATCTTGGCCAGGAGGAGAGAAATGTCTTGATCGAATATCTTAAAGA	343
Db	284	AGCTCTTCTTGTATCTTGGCCAGGAGGAGAGAAATGTCTTGATCGAATATCTTAAAGA	343
QY	344	ATGAACTGGAACAATGTCAAGAGCCAGCTTTTCCAGAAAGCAAGAGAAACGAGACAGCC	403
Db	344	ATGAACTGGAACAATGTCAAGAGCCAGCTTTTCCAGAAAGCAAGAGAAACGAGACAGCC	403
QY	404	AGGTCATCATGACACACTCTGGGGATACGGTGGAAAGCAACATGCTACGTATCTC	463
Db	404	AGGTCATCATGACACACTCTGGGGATACGGTGGAAAGCAACATGCTACGTATCTC	463
QY	464	TGCACGAGCCCTTGGGCAAGGCCAGAGATGCTGTGCTCCACACTGAAAAAAGCAGATTAAGT	523
Db	464	TGCACGAGCCCTTGGGCAAGGCCAGAGATGCTGTGCTCCACACTGAAAAAAGCAGATTAAGT	523
QY	524	ACTTAAAGCAGCAGCAGAGATGTAACCAACAACAGCAAGAGAGGCGCGGCTCAGCA	583
Db	524	ACTTAAAGCAGCAGCAGAGATGTAACCAACAACAGCAAGAGAGGCGCGGCTCAGCA	583
QY	584	GCAACATGAACACCATGAGAGAGATTTAGCTTCACTCCAGAGCCAGCTCCCTGAGGTGG	643
Db	584	GCAACATGAACACCATGAGAGAGATTTAGCTTCACTCCAGAGCCAGCTCCCTGAGGTGG	643
QY	644	AGGAGATGATCCGAGACATGGGGTGGGAGACAGTCAGCGGTGGAACAGTGGCTGTGTACT	703
Db	644	AGGAGATGATCCGAGACATGGGGTGGGAGACAGTCAGCGGTGGAACAGTGGCTGTGTACT	703
QY	704	GTGTCTCTCTCAAGAAAGATGACGAGATCTAAAAAGGCGCAGGAAAGCCTCAGGGGAGG	763
Db	704	GTGTCTCTCTCAAGAAAGATGACGAGATCTAAAAAGGCGCAGGAAAGCCTCAGGGGAGG	763
QY	764	TGGCGTGAACACTGAGAGAGATTTGTTTCCCTCCACAAGCAAGTTGCAGACGTCTACT	823
Db	764	TGGCGTGAACACTGAGAGAGATTTGTTTCCCTCCACAAGCAAGTTGCAGACGTCTACT	823
QY	824	CTGAATTTGATTCAGAGCCAAAGTTAAGCTAGAGCCCAAGAGACTTTCACAGTGTG	883
Db	824	CTGAATTTGATTCAGAGCCAAAGTTAAGCTAGAGCCCAAGAGACTTTCACAGTGTG	883
QY	884	ACAAGAAATCATGAGCCCTGAAAAAAGACTTAACGATGTCAGAGAAACCTTGAACTGC	943
Db	884	ACAAGAAATCATGAGCCCTGAAAAAAGACTTAACGATGTCAGAGAAACCTTGAACTGC	943
QY	944	CACCACTGGCCAGTGAAGACTGTCCACGCGCTGGTTTAAAGAGCCCGACGCTGTGGAGG	1003
Db	944	CACCACTGGCCAGTGAAGACTGTCCACGCGCTGGTTTAAAGAGCCCGACGCTGTGGAGG	1003
QY	1004	TGAATCTGAAGCTCGCCGGGCACTTCCTCGTGATGATTTGAATCTCAATGATACCTTGG	1063
Db	1004	TGAATCTGAAGCTCGCCGGGCACTTCCTCGTGATGATTTGAATCTCAATGATACCTTGG	1063
QY	1064	ATGTGGATACTCCCCCAAGCCCGGCTCTCCAGCTCCAGCATGGTTACTACGAAGAACTTT	1123
Db	1064	ATGTGGATACTCCCCCAAGCCCGGCTCTCCAGCTCCAGCATGGTTACTACGAAGAACTTT	1123
QY	1124	GCCTTGAAGAAGTCACACTTCCCAATTCAGGATGTCTCCCAAGAGATATGCAAAAGGCCCA	1183
Db	1124	GCCTTGAAGAAGTCACACTTCCCAATTCAGGATGTCTCCCAAGAGATATGCAAAAGGCCCA	1183
QY	1184	GGAAGAGATCCAGCTCACTGGGGTGGCCAGACTCTGCAAGAGAGGCCAGATGAGAGAAC	1243
Db	1184	GGAAGAGATCCAGCTCACTGGGGTGGCCAGACTCTGCAAGAGAGGCCAGATGAGAGAAC	1243
QY	1244	TGGTTGTGCTTCCTCTATTTTTCGCGAATGCACTCTAAGGCCAGAAACACGCCAANA	1303
Db	1244	TGGTTGTGCTTCCTCTATTTTTCGCGAATGCACTCTAAGGCCAGAAACACGCCAANA	1303
QY	1304	GGCCCAAGTCAAGAGTCTCTTTGCAAGCAAAAGATGTGTGAAGACAGGCTTGATGGGCTCG	1363
Db	1304	GGCCCAAGTCAAGAGTCTCTTTGCAAGCAAAAGATGTGTGAAGACAGGCTTGATGGGCTCG	1363

OY	1364	GGGCGCGGCAAAATTTCAATCCAGACCTAGACAACAAGTCAATGAATCCGCCCATTCGCCTGTAA	1423
OY	1364	GGGCGCGGCAAAATTTCAATCCAGACCTAGACAACAAGTCAATGAATCCGCCCATTCGCCTGTAA	1423
OY	1424	AGCCCCAAGACCAGAAGTTTAAGCAGAGGCTGAGGGTGTAAGAACCCCTGCCTTCTCTTCCAGG	1483
DB	1424	AGCCCCAAGACCAGAAGTTTAAGCAGAGGCTGAGGGTGTAAGAACCCCTGCCTTCTCTTCCAGG	1483
OY	1484	CCAAGCTGGACACCTTTCCTGCTGCTGTGAACAAGTAGTGAAGTGCACCAATGGCCAGACACA	1543
DB	1484	CCAAGCTGGACACCTTTCCTGCTGCTGTGAACAAGTAGTGAAGTGCACCAATGGCCAGACACA	1543
OY	1544	TGCGTGCACACTGTAGTGAAGAGCTGTCCAGGACGGG - TTTGGTGCAGAGGCCCTACT	1601
DB	1544	TGCGTGCACACTGTAGTGAAGAGCTGTCCAGGACGGG - TTTGGTGCAGAGGCCCTACT	1601
OY	1602	TTTCGGGACAGCCTTAGAGTGTAAAGGCGACAGCAAACAGGTGAGGGTGAAGTGTACACCCAG	1661
DB	1604	TTTCGGGACAGCCTTAGAGTGTAAAGGCGACAGCAAACAGGTGAGGGTGAAGTGTACACCCAG	1663
OY	1662	AGACTGCTCTTCCCTGACCTCACCTGACCCACTCTACGACTGGAGAGTGCATGACACAG	1721
DB	1664	AGACTGCTCTTCCCTGACCTCACCTGACCCACTCTACGACTGGAGAGTGCATGACACAG	1723
OY	1722	CCCCTGATCCCTGTAGCAGCAGGTCCTGCT - CTGTGGCAGAGCTCTTGTATTATGCCATGAT	1780
DB	1724	CCCCTGATCCCTGTAGCAGCAGGTCCTGCTCTTGTGGCAGAGCTCTTGTATTATGCCATGAT	1783
OY	1781	CAGATGTGGTCAAGACTCTTCTGGGCGCTTGGAGACCAAGCGGTCACTTGTTGACATGCTGCT	1840
DB	1784	CAGATGTGGTCAAGACTCTTCTGGGCGCTTGGAGACCAAGCGGTCACTTGTTGACATGCTGCT	1843
OY	1841	GGACCAAGAGTGCCTTAGAGCATCTCAGGCGACGCTCAGGCCCACTTCTACCTGCTTTGAC	1900
DB	1844	GGACCAAGAGTGCCTTAGAGCATCTCAGGCGACGCTCAGGCCCACTTCTACCTGCTTTGAC	1903
OY	1901	TTGCTTTCTTA - GCATAGCCCTGGGGCCAAAGCAGGGTGGGGGAATGAGAGATAGACATGGGATGT	1959
DB	1904	TTGCTTTCTTAAGCATAGCTGAGCTGGGGCCAAAGCAGGGTGGGGGAATGAGAGATAG - CATGGGATGT	1962
OY	1960	ATGCAAGAGATGGAAGATTTTCCCAAAAAAAAAAAAAAAAAAAAAA 2007	
DB	1963	ATGCAAGAGATGGAAGATTTTCTATGTAAAATAATTAATAAAAAAAAAA 2010	
RESULT	6		
LOCUS	BCO17374	2024 bp mRNA linear ROD 14-NOV-2001	
DEFINITION	Mus musculus, TRAF-interacting protein, clone MGC:13919		
ACCESSION	BCO17374		
VERSION	BCO17374.1 GI:16924208		
KEYWORDS	MGC.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 2024)		
JOURNAL	Straussberg,R. Direct Submission Submitted (13-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: amadan@systemsbio.org		

Query Match	Best Local Similarity	74.9%	Score 1087	DB 10	Length 2024
Matches 1492	Conservative	0	Mismatches 440	Indels 61	Gaps 8
26	AGCGGAGCGGGGCGCTTACGAAAGCCGAGACCTGTAGCAGTTTCTTTGGCTGCTGGGCC	85			
26	AGCGGCGCGCGTTCGGCGCACCACCAACTGTGCTGTGCGTG-GCAAGCTGGTTCCCGGGGCT	84			
86	CTTGTAGTCCAGCCATCATGCGCTATCCGTCGTGTGACATATGCTGCGACTTCTTGG	145			
85	GCTTGTAGTCCAGCCATCATGCGCTATCCGTCGTGTGACATATGCTGCGACTTCTTGG	144			
146	ATCACTTCCCGGAGCTGGCGCCGATCATCTGCGAGCCAGCCCTTCCACTTGTGAGTCCCTAA	205			
145	ATCACTTCCCGGAGCTGGCGCCGATCATCTGCGAGCCAGCCCTTCCACTTGTGAGTCCCTAA	204			
206	TTCACTTCTTGTGAGCAGCACCAGTGGAGCTGGCCAGCTGCCGAATCCAGTTGGCA	265			
205	TTCATGTGTTGTGAGCAGCACCAGTGGAGCTGGCCAGCTGCCGAATCCAGTTGGCA	264			
266	AAAGAACATATATCAATTAAGCTCTTGTGATCTTGTGCGCAGAGAGAGAGAAATGCTTGG	325			
265	AAAGAACATATATTAACAACCTTCTTGTGATCTTGTGCGCAGAGAGAGAGAAATGCTTGG	324			
326	ATCGAGATTTCTTAAAGAAATGACATGACATGTTCAGAGCCAGCTTTCAGAAAGACA	385			
325	ATCGAGATTTCTTAAAGAAATGACATGACATGTTCAGAGCCAGCTTTCAGAAAGACA	384			
386	AGGAGAAACGAGACAGCGAGTCAATGACATCTGCGGGAATACGCTGGAAGAACCA	445			
385	GGGAGAAACGAGACAGCGAGTCAATGACATCTGCGGGAATACGCTGGAAGAACCA	444			
446	ATGCTACTGTGTATCTTGTGACAGAGGCTTGGGAGAGGCGGAGATGCTGTGCTCCACAC	505			
445	ATGCTACTGTGTATCTTGTGACAGAGGCTTGGGAGAGGCGGAGATGCTGTGCTCCACAC	504			



QY	506	TGAAAAAGCAATGAATACTTACTTAGAGAGACAGAGATGAGACCAAAACAAGCAAGAG	565
Db	505	TGAAAAAACAATGTAAGTCTCTGGAGCAGGGGAGAGATGAGACAAACAAGCTCGGAGG	564
QY	566	AGGCGGGCGGGCTCAGAGAGCAAGTGAAGACCAATGAGCGAGATTGAGCTTCACTCACA	625
Db	565	AGGCCCAACCGACTCAAGTGCAGATGTAAAAACCATGGAGCAATTGAGCTCTACTCAGA	624
QY	626	GCCAGCTCCCTGAGGTGGAGAGATGATCCGAGACATGGGTGTGGGACAGTCAGCGGTG	685
Db	625	GCCAGCGGTCTGAGGTGAGAGAGATGATTCGAGACATGGGTGTGGGACAGTCAGCGGTG	684
QY	686	AACAGCTGCGCTGTACTGTGTGTCTTCAAGAAAGATACGAAATCTTAAAAAGCGAC	745
Db	685	AGCAGCTGCGCTGTACTGTGTGTCTTCAAGAAAGATGTGAATCTGAAGGAAGCTC	744
QY	746	GGAAGGCTCAGGGGAGGTGGCTGACACAGCTGAGGAAGGTTGTTTCTCAGAGCA	805
Db	745	GGAAGGCTCAGGGGAGGTGGCTGACAGGTTTGAAGAAGGATTTGGTGTCTTCAGAGTA	804
QY	806	AGTTGCAGACAGTCTACTCTGAATTGGATAGGCCCAAGTTAGCACTGAAGTCAGCCACA	865
Db	805	AGTTGAGAGACTCTCAACACTGAGCTGGATCAGGCCAAGTTAGACTGAGGTACGCCACA	864
QY	866	AGGACTTACAGATGCTGTGACAGAGAAATCATGAGCCTGAAAAAAGACTAACGATCTGC	925
Db	865	AGGACTTACAAAGTGTGACAGAGATCATGAGCCTAAGAAAGAGCTGATGATCTCTCC	924
QY	926	AGGAACCTTGAACCTGCGCACAGTGGCCAGTGAAGCTGTGACCGCGCTGTGTTTAGAA	985
Db	925	AGGGAACCTTGAAGCTCTGCTCCGCGCACATGAGAGGTCACCGCTGTGTTTGAAGA	984
QY	986	GCCAGCGCCCTGTGGAGGTG---AATGTGAAGTCCGCGCGGCGCATCTCCGATGATA	1045
Db	985	GCCAGCGCCCTGTGGAGATGTGAACCCGAGGCTTCCACAGGCCACCTTGTGGATGAGA	1044
QY	1043	TTGATCTCAATGCTACCTTTGATGTGATTACTCCCGACGCGCGCCCTCAGCTCCACG	1107
Db	1045	TTGATCTCAATTAACCACTTTGATGTAAATACCCCTCCACACCCAGACTCTGGGTCCAC	1106
QY	1103	ATGCTTACTACGAAAAATTGGCTTAGAGAAATCACACTGCCCAATTCAAGATGTCCCA	1166
Db	1105	ATTGCTCTCCCAACAACCTGTGCTGTGAGAGGGCAGCGCTCTCCATGCAAGATGTCTCA	1165
QY	1163	AGAAGATATGCAAAAGGCCCAAGAGAGAGTCCAGAGCTCAGTGGGTGGCGAGAGCTGG	1222
Db	1165	AGAAAGTGCACAAAGTCTCCAGCCGGAGTCCAGAGCTCTACGTGGGTGGCGAGCGATGG	1222
QY	1223	CAGGAGCCAGATGAGAGAACTGGTTGGTCTCTCCCTATTTTGTCCGGAATGCCATCC	1283
Db	1225	TAGGAGAGCTGATGAGGAACGTGGCTGTGCTCTCCCTCTTCAATCCGAATGCTGTCC	1282
QY	1283	TAGGCCAGAAAACAGCCCAAAAAGCCCAAGTCAAGTCTCTTGCACAAAGATGTGTAA	1344
Db	1285	TGGGTCAAGAAAACAGCCCAACAGGACCCACACAGCAATCCCTGTGCACACAGATGTGTAA	1344
QY	1343	GGACAGGCTTGATGGGGCTGGGTGGCCGGCAAAATTAATCCAGCCTACTGACACGTGA	1405
Db	1345	GAATAGGCTTGAATGGGCTTGGAGAGCAAAATTAATCAATCCAGCCTTAGGGACAAACCA	1404
QY	1403	TGATCCCGCCCTTACCTCTGTTTAAGCCCAAGCAAGGTTATGACAGAGGTGAGGTTAAGA	1465
Db	1405	TTATCCGACACGTGCTGTGTTAATGTTCAAGGCCAAGAGTAATACGAAAGATGAGATTAAGA	1466
QY	1463	CCGTGCTTCTCTCTCCAGGCCAAGCTGACACCTTCTGTGTCGTGAGAACAGTGAAG	1522
Db	1465	CTGTGAGTTCCTCTCCAGGCCAAGCTGATACCTTTATGTGAC-----	1521
QY	1523	TCTGACCAATGGCCAGACACATGCTCTCAACTGTGAGTGAAGAGACTGTCCA--GGCAG	1588
Db	1512	--TGAGAGGTACCAAGATGATGTTTCAATTTATGTGGCCCAACACCTGGCTTAACCGGAAG	1566

OY		1581	GTTTCTGACAGACGCCCTACTTTTGSGAACCAAGCCTGGAGGTGTAAAGGCAGACAACAGAGT	1640
Db	1570	TGTTTTGGAAGAATGGCTTCCTTTGGACC-----		1598
OY	1641	GAGGCTGAGTGTGACACCAGAGACTGCTTTCCTCGCCCTCACCTGCCCATCTAAGC	1700	
Db	1599	-AGTCCAGAGAGATAGCCCCAGAAAAACACATTCTCTGTTTACTCTGCGCCCTGCACAC--AC	1655	
OY	1701	ACTGGACCTGACATGACCAAGCCCACCTGATTCCTGTCAAGACAGTCTGCG-TCGTHTGCGAG	1759	
Db	1656	ACTGGAGACCAATGACAGTAGTACTGTCTCGATACACAGAGGCGCTACTTCAGATTGGAG	1715	
OY	1760	GCTCTGTTTAATAGCATGATCAATGTGTGTGCACACTTTTCTGGCGCTGSAGACACAGG	1819	
Db	1716	GATTTTGGCTTATAGTATACACCAAGGTGGCGGTGACCTCTTTGTTTTATAAGACAGGG	1775	
OY	1820	TCACTTGTGACTGCTCTGTGGACCAG-----AGTGGTTAAGCATATCTCAGGACAGCCTC	1874	
Db	1776	TCACATTGATCTTAAGTGGATGGAGGTCTCGAGAGATCTTATGCAGGCTGTAGACACCTTG	1835	
OY	1875	A GCCCAAGCTTCTACCTCCCTTTGACTTGCTTTCAGCATAGCATCGGCCCAAGAGGGTGG	1934	
Db	1836	C GCTTGAACCTCGCTGCCCTCCACACTTATTTGCTTGAAAATATAGGGGATGGAGTGGTATA	1895	
OY	1935	GGAATGAGAGATAGACATGAGGATGTATGAGAGAGATGAGACATTTTCCCGAAAAAAAA	1994	
Db	1896	G GGAAGAGCTTGGGGAAGTTTCTGTGTAAATAAAGACATCTTTTCTTCAAASAAAAA	1955	
OY	1995	AAAAAAAAAAAAA 2007		
Db	1956	AAAAAAAAAAAAA 1968		
RESULT	7			
MWU77844				
LOCUS	MWU77844	1981 bp	mRNA	linear ROD 25-Apr-1997
DEFINITION	Mus musculus mTRIP (mTRIP) mRNA, complete cds.			
ACCESSION	U77844			
VERSION	U77844.1 GI:2039305			
KEYWORDS	.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus-			
TITLE	Lee,S.Y., Lee,S.Y. and Choi,Y. TRAF-interacting protein (TRIP): a novel component of the tumor necrosis factor receptor (TNF)- and C3D1-TRAF signaling complexes that inhibits TRAF2-mediated NF-kappaB activation J. Exp. Med. 185 (7), 1275-1285 (1997)			
JOURNAL	97258620			
MEDLINE	2 (bases 1 to 1981)			
REFERENCE	Lee,S.Y., Lee,S.Y. and Choi,Y.			
AUTHORS	Direct Submission			
JOURNAL	Submitted (10-NOV-1996) Immunology, The Rockefeller University, 1230 York Avenue Box 295, New York, NY 10021, USA			
FEATURES	Location/Qualifiers			
Source	1..1981			
	/organism="Mus musculus"			
gene	/db_xref="taxon:10090"			
	1..1981			
	/gene="mTRIP"			
CDS	118..1530			
	/gene="mTRIP"			
	/codon_start=1			
	/product="mTRIP"			
	/protein_id="AAB52994.1"			
	/db_xref="GI:2039306"			
	/translation="MPTLSICTLASEPFDDHSRDVAATHGHGTHFLQCLIQMFAPSR TCPCKRIYGVKKITINKLFPIADAEENVIDAEPLKKEIDSVKAQLSKRREKRDSA IIDTLRLTEERNATVESLOALNALKMAELSLTKQKFLEQRODETFKQAREEHRIK CEMTDRKDIEILLQSSEWEEMIRDMVGQAQSAEQLAVCVSLIKYEIMLKARAT GLNDLRDKDIIVSSRSKLKITLNLELDQAKLETLSAQCDLOSADDELISLKKKIILIG			

BASE COUNT	532 a	489 c	534 g	426 t
ORIGIN	TSLPRTNETVSLVPE\$APVEMMRLHQPEGEIDNTFEDVNTPTQTS\$GQ HCLPKLCLEBAR\$PMONLVKRVK\$PESQLG\$ORCVGELDEELAG\$FPLTIN\$ AVLQ\$KQ\$NRTTASR\$SDVVRIG\$FGLG\$RTFIO\$PRDITTIIRPVV\$SKAS\$KOR\$ VRKTSAS\$PKIDTFLCQ"			

Query Match	53.9%;	Score 1081;	DB 10;	Length 1981;
Best Local Similarity	74.6%;	Pred. No. 1.6e-291;		
Matches 1504; Conservative	0;	Mismatches 450;	Indels 63;	Gaps 9

QY	1	GTGGGGTGGACGGAATTTGAAGCAACGGGAGGGGGC--GCCTCTCGAAGCGGAGCT	58
Db	15	GTGGGGTGGACGGAATTTGAAGCAACGGGAGGGGGTTCCACCAACTGTGTCT	74
QY	59	GTACACATTTCTTTGGCTGCTGGGCCCCCTTGAATCCAGCCATCATCTCGTCTC	118
Db	75	GTCTCTG-GCAGCTGGTTCCTCTGGGCTCTTGAATCTCAGCCATCATCTCTCTCTC	111
QY	119	TGTGCACATATGCTCGGACTTCTTCGATCACTCCCGCAGCGTGGCGGCATCCATCGC	178
Db	134	TGTGCACATATGCTCGGACTTCTTCGATCACTCCCGTGAAGTGGCTGCCATCCATCGT	193
QY	179	GCCACACCTTCCACTTGGCAGTGGCTTAATTCACTCTTTGAGACGACCAAGTGGACCT	238
Db	194	GCCACACTTTTCACCTCGCAAGCCTTAATCCAGTGGTTGAGACGACCAAGTGGACCT	253
QY	239	GCCACACGTGCGCAATCCAGGTTGGCAAAAAGAACATTATCAATTAACCTCTCTTCATC	298
Db	254	GCCACACGTATAGATTCAGGTTGGCAAAAAGACATTATTAACCAACTTTCTTTGACC	313
QY	299	TTGCCACAGAGGAGAGAGATGCTTGGATCGAGATTTCTTAAGATGAACCTGGACATG	358
Db	314	TCGGCCACGGAAGAGAGATGCTTGGATCGAGATTTCTTAAGATGAACCTGGACAGCG	373
QY	359	TCAGAGCCACCTTCCACAGAAACAGGGAAGAAAGACAGACGCGACGTATCATGACACA	418
Db	374	TCAAAGCTCAGCTTTCCACAGAAACAGGGAAGAAAGACAGACGCGACATTCAGACA	433
QY	419	CTCTGGGGATACGCTGGAAGAACGCAATGCTACTGTGATCTCTTGACAGAGCCTTGG	478
Db	434	CTTACGGGACACCTCGAAGAAAGCAATGCTACGTHGAGTCCCTTACAGAACGCTTAA	493
QY	479	GCAAGCGCGAGATGCTGTGCTCCCACTGAAAAAGCAGATGAAGTCTT7AGACGACAGC	538
Db	494	ACAAAGCAGAGATGCTGTGCTCCACCTGAAAAAGATGAAGTCTTGACAGAGCGGC	553
QY	539	AGGATGAGACCAACACAGCACAAAGAGAGGCGGCGCTCAGAGGCAAGATGAAGCCA	598
Db	554	AGGATGAGACCAACCAAGCTCGGAGGAGGCGCCAGCTCAAGTGCMAATGAAAAACA	613
QY	599	TGGAGCAGATTTGAGCTTCTACTCTCCAGACCAAGCTCCCTGAGGTGGAGAGATCCGAG	658
Db	614	TGGAGCAAAATTGAGCTCTCTACTCCAGAGCCACGCGTTCTGAGGTGGAGAGATGATTCGAG	673
QY	659	ACATGGGTTGGGACAGCTCAGCGGTGGAACAGCGGTGTGTTCTGTCTCTCAAG	718
Db	674	ACATGGGTTGGGACAGCTCAGCGGTGGAACAGCGGTGTGTTCTGTCTCTCTCAAGA	733
QY	719	AAGAGTACGAGAACTCTAAAGAGGCAAGGAGGCTTAGGGGAGGTGGCTGACAAAGCTGA	778
Db	734	AAGAGTATGAGAACTCTGAAGAGAGCTCGAAGGCGCACAGGGGAAGGCTGACAGGTTGA	793
QY	779	GGAAGATTTTCTTCTCTCCAGAAAGCAAGTTGCAGACAGTCTACTCTGAATTGGATCAG	838
Db	794	AGAAGATTTTGGTCTCTCTTAGAGGCAAGTTGAAGACTTCAACACAGAGCTTGATCAG	853
QY	839	CCAAGTTAGAACTCAATCTCAGCCAGAGAGACTTACAGATGCTGTCGACAGGAATCATGA	898
Db	854	CCAAGTTAGAACTCATGAGTCAAGCCAGAGAGACTTACAAAGTGTCTGACACGAGATCACGA	913
QY	899	GCCTGAAAAAGAGCTAACGATGCTGCAGGAAACCTTGAACCTGCCACAGCTGGCCAGTG	958

Db	914	GCCTAAGAAAGAAAGCTATGATATCTCCAGGGAACCTTGACCTTCGCGCTACCAATG	973
Qy	959	AGACTGTGACACCGCTGTTTATAGAGAGCCACCCCTGTGGA--GGTAATCTGAAGC	10151
Db	974	AGACGGTCAGCGCGCTGTTTTTGAAGAGCCACCCCTGTGGAGATGATGAACCCGAGGC	10333
Qy	1016	TCGCGCGGCACTCTTCCGTGATGATGTGATCTCAATGCTACCTTGTGATGTACTC	10757
Db	1034	TTCAACCAAGCCACCTTGGTATGAGATTGATCTCAATACCACTTGTGATTAATACC	10933
Qy	1076	CCCCAGCGCCGCTCCAGCTCCACAGATGTTACTACGAAAAACCTTCCCTAGAGAGT	11333
Db	1094	CTCCAAACCCAGACCTCTGCGCTCCAGCATGGCTCTCCCAAGAGCGTGCCTGAGAGGG	11533
Qy	1136	CACACTCCCCAATTCAGAGATGTCCCCAAGAAGATATGCAAAAGGCCCAAGAGAGTCCC	11993
Db	1154	CACGCTCTCCCATGACAGAAATGTCTCTCAAGAAGTGCACAAAGTCTCAAGCCGAGATCC	12133
Qy	1196	AGCTCTCACTGGGTGGCCAGAGCTGTGCAGAGAGCCAGATGAGAACTGTTTGGTGCCT	12553
Db	1214	AGCTCTCACTGGGTGGCCAGAGCTGTGTAGAGAGCTAATAGGAACCTGCTGGTGCCT	12733
Qy	1256	TCGCTATTTTTTGTCCGGAATGCCATCTAGGCCAGAAACAGCCCAAAAGGCCAGTCAAG	13153
Db	1274	TCGCTCTCTTATCTCGGAATGTCTCTCTGGTGTAGAAACGCCCAACAGACCAAGCAG	13333
Qy	1316	AGTCTCTTTCAGCAAGATGTGTGAAGCAGAGCTTCGATGGGTCGGTGGCCGAGCAA	13757
Db	1334	AATCCCAAGAGACGACGATGTGTGAAGATAGCTTTATGTGGCTTGGAGGACACACAA	13933
Qy	1376	AATTCATTCACAGCTACTGACACAGTCATGATCCGCCATTTGCTGTAAAGCCAAAGCCA	14353
Db	1394	AATTCATTCACAGCTTAGAGGACCAACCATTAATCCGACCAATGCTGTAAAGTCAAGGCCA	14533
Qy	1436	AGGTAAAGCAGAGGGGTGAGGGTGAACACCGTCTCTCTTCCAGGCAAGCTGGACAA	14953
Db	1454	AGAGTAAGCAGAAAGTGAATTAAGACTGTGATTTCTCCTCCAGGCCAAGCTGGAGTA	15133
Qy	1496	CTTCTCTGTGGTGTGTGAACAGTGTGATCTGACCAATATGGCCAGACACATGCTGCACATT	15553
Db	1514	CTTCTTATATGCAG-----TGAACGGTGAACAGAGATGTTTGCATTA	15553
Qy	1556	GTAAGTCAAGACGTGTCA--GGCAGGGTTTGTGACAGAGCCCTACTTTCGGAGCCAGC	16133
Db	1559	GTTGGGCAAGAACCTGGCTAACCGGAAGTGTTTTGGAAATGATGGCTCTCTTGAGCC---	16143
Qy	1614	CTGAGGTGTAAAGGCAACAAGAGTGAAGGTGATGTGAACCCACAGACTGCTCTTC	16733
Db	1615	-----AOTCCAAAGAGATGCCCAAGAAAACACACTTC	16463
Qy	1674	CTGCGCTCACCCCTGCCCCACTCTTACGACGTGGAGGTGATGATGACAGGCCACTGATCTT	17333
Db	1647	CTGTGTTTCACTGGGCCCTGTGACC--ACTGGAGAACCATATGACCAATTACTTTCGG	17043
Qy	1734	GTCAGCAGGTCTGTC--TCTGTTGCCAGGCTCTTGTATTATAGCCATGATCAGATGTGTCNA	17923
Db	1705	ATCAGCAGGGCTACTTCCAGTTCAGTGGATTTGCTTATAGTCACMACAGGTGTGGCTG	17643
Qy	1793	GACTCTTTTGGGCGCTGGAGACCAAGGTCACTTGTGACTGTCTGTGGACCAAG-----	18473
Db	1765	GACTCTTTTGTATTATAGAAAGGTGCACATTCGCTAAGTGAATGGAGTGGCTGGA	18243
Qy	1848	AGTCTTGAAGACATTCAGGCACTCTCAGGCCCAAGCTTTTACCTGCTTGTGACTGTTC	19073
Db	1825	GGATCTTATGCAAGCTGGAGAGACCTGCGCTTGAATCTCTGCTCTCAGCTTATATGC	18843
Qy	1908	TAGCATAGCTGGGCGCAAGCAGAGGTGGGAATGAGAGATAGACATGAGGATGTATGAGAG	19673
Db	1885	TTGAAATTAATGGGTGAGGTGTGATTAAGGAAAGTGTGGGAAGTTTCTGTATAAATA	19443
Qy	1968	GATGGAAGATTTTCCGAAAAAAGAAAAAAGAAAAA 2004	

Db 1945 AAAAGGATCTTTCTTCAAAAAAAAAAAAAAAAA 1981

RESULT 8  
LOCUS BC006929 3694 bp mRNA linear ROD 12-JUL-2001  
DEFINITION Mus musculus, TRAF-interacting protein, clone MGC:11463  
IMAGE:3156502, mRNA, complete cds.  
ACCESSION BC006929  
VERSION BC006929.1 GI:13905261  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Strausberg, R.  
REFERENCE 1 (bases 1 to 3694)  
AUTHORS Direct Submission  
TITLE Submitted (27-APR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: Jeffrey Green M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [villalona@bcm.tmc.edu](mailto:villalona@bcm.tmc.edu)  
Villalon, D.K., Luna, R.A., Hale, S.M., Huljk, S., Lu, X., Garcia,  
A.M., Holloway, M., Teiford, B., Hodgson, A., Bouck, J., Yu, W.,  
Muzny, D.M., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAC Plate: 16 Row: a Column: 23  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 2039305.  
Location/Qualifiers  
1. 3694  
/organism="Mus musculus"  
/db\_xref="LOCUSID:22036"  
/db\_xref="taxon:10090"  
/clone="MGC:11463 IMAGE:3156502"  
/tissue\_type="Mammary tumor, C3(1)-Tag model, infiltrating  
ductal carcinoma, 5 month old virgin mouse."  
/clone.lib="MCI.GCAP\_Mam6"  
/lab\_host="DH10B"  
/note="Vector: PCMV-SPORT6"  
277. .948  
/codon\_start=1  
/product="TRAF-interacting protein"  
/protein\_id="AAH06929.1"  
/db\_xref="GI:13905262"  
/translation="MPRALCTICSPFFHSRDVAIIHCHTFHLOCLIQWETTPASR  
ACPOCRIVQCKKTIINKLFEDLAQEEENVLDADFLKNEIDSVAAQISOQKREKRSQA  
IIDTLRDTLEERNAIVESLONALNKAEMLCSTLKQMKPLEQRODETKQAREDAHRIK  
CKMKTMEQIEILLQSORSSEVEEMIRDMVGQSAVEQLAVYCVSLKCVSSGVDLPSSF  
VMQIL"  
BASE COUNT 858 a 951 c 976 g 909 t  
ORIGIN

Query Match 26.0%; Score 521.6; DB 10; Length 3694;  
Best Local Similarity 87.8%; Pred. No. 1,1e-134;  
Matches 569; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

72 TGCGTGGCTGGGCGCTTGAGTCAGCCATCATGCCATTCGTCGTGCACTATCTG 131  
|||||

Db 246 TGGTTCCTGGGCTGCTTGAAGTCAGCCATCATGCCATTCGCGCTCTGTGCACTATCTG 305

QY 132 CTCGCACTCTTGATCATCTCCGCGACGTCGGCGCATCGACATGGGCGCCACCTTCCA 191  
|||||  
Db 306 CTCGCACTCTTGATCATCTCCGCGACGTCGGCGCATCGACATGGGCGCCACCTTCCA 365  
|||||

QY 192 CTTCAGTGGCTTATTCAGTCCTTTGGAGACAGCACAAGTCGGACCTGCCCCAGAGTCCG 251  
|||||  
Db 366 TCTGCATGGCTTATTCAGTCGTTTGGAGACAGCACAAGTCGGACCTGCCCCAGAGTCCG 425  
|||||

QY 252 AATTCAGGTTGGCAAAAGAACATTAATTAAGCTCTTTTGATCTTCCAGAGAGGA 311  
|||||  
Db 426 AATTCAGGTTGGCAAAAGAACATTAATTAAGCTCTTTTGATCTTCCAGAGAGGA 485  
|||||

QY 312 GGAGAAATGCTTGATGAGCAATTTCTTAAGAAATGAAGTGAACATGTCAGAGCCAGCT 371  
|||||  
Db 486 GGAGAAATGCTTGATGAGCAATTTCTTAAGAAATGAAGTGAACATGTCAGAGCCAGCT 545  
|||||

QY 372 TTCCAGAAAGACAGAGGAGAGAGAGACAGCAGTCATCATGACACTCTGGCGGATAC 431  
|||||  
Db 546 TTCCAGAAAGACAGAGGAGAGAGAGACAGCAGTCATCATGACACTCTGGCGGACAC 605  
|||||

QY 432 GCTGGAAGAACGCAATGCTACTGTGTATCTCTGACAGAGCCTTGGGCAAGCCGAGAT 491  
|||||  
Db 606 CTTGGAAGAACGCAATGCTACTGTGTATCTCTGACAGAGCCTTAAACAGGACGAGAT 665  
|||||

QY 492 GCTGTGCTCCACACTGAAAAAGAGATGAAGTACTTGACAGCAGCAGAGATGAGACCA 551  
|||||  
Db 666 GCTGTGCTCCACACTGAAAAAGAGATGAAGTACTTGACAGCAGCAGAGATGAGACCA 725  
|||||

QY 552 ACAAGCACAAGAGAGGCGGCGGCTCAGAGCAGCAAGATGAAGACATGAGAGCAGATTTGA 611  
|||||  
Db 726 ACAAGCACAAGAGAGGCGGCGGCTCAGAGCAGCAAGATGAAGATGAAGACATGAGAGCAGATTTGA 785  
|||||

QY 612 GCTTCTACTCCAGACCAAGTCTCCTGAGGTGGAGAGATATCCGAGACATGGGTGGG 671  
|||||  
Db 786 GCTTCTACTCCAGACCAAGTCTCCTGAGGTGGAGAGATATTCGACATGAGGTGGG 845  
|||||

QY 672 ACAGTCAGGCGTGGAGACAGCTGGCTGTACTGTGTCTCTCAAGAA 719  
|||||  
Db 846 ACAGTCAGGCGTGGAGACAGCTGGCTGTACTGTGTCTCTCAAGAA 893  
|||||

RESULT 9  
LOCUS HS465N24 129747 bp DNA linear PRI 23-NOV-1999  
DEFINITION Human DNA sequence from clone 465N24 on chromosome 1p35.1-36.13.  
Contains two novel genes, ESTs, GSSs and Cpg Islands, complete  
sequence.  
ACCESSION AL031432  
VERSION AL031432.1 GI:4375969  
KEYWORDS HNG; Cpg Island.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 129747)  
AUTHORS Wilson, S.  
TITLE Direct Submission  
JOURNAL Submitted (11-FEB-1999) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
On Mar 7, 1999 this sequence version replaced gi:4176479.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence is the entire insert of clone 465N24. This sequence  
has been finished according to sequence map criteria as follows. An  
attempt is made to resolve all sequencing problems, such as  
compressions and repeats, but not necessarily within known

annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.  
This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>  
465N24 is from the library RPC13 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/VECTOR>.  
pcrPAC.

## FEATURES

Source  
1. 129747  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="p35.1-36.13"  
/clone="RP3-465N24"  
/clone\_lib="RPC1-3"  
540. 753  
/note="L2 repeat: matches 2255. .2461 of consensus"  
975. 1303  
/note="AluI repeat: matches 1. .311 of consensus"  
1320. 1547  
/note="MIR repeat: matches 29. .262 of consensus"  
2061. 2268  
/note="MIR repeat: matches 100. .262 of consensus"  
2356. 2518  
/note="MER5A repeat: matches 5. .164 of consensus"  
2519. 2827  
/note="AluIb8 repeat: matches 1. .318 of consensus"  
3245. 3360  
/note="MIR repeat: matches 20. .139 of consensus"  
3427. 3485  
/note="L2 repeat: matches 2688. .2746 of consensus"  
3444. 3676  
/note="MIR repeat: matches 8. .249 of consensus"  
3690. 3856  
/note="LTR41 repeat: matches 20. .182 of consensus"  
4187. 4494  
/note="AluSg repeat: matches 1. .308 of consensus"  
4846. 5143  
/note="AluS repeat: matches 1. .296 of consensus"  
5287. 5340  
/note="L1MD1 repeat: matches 5345. .5398 of consensus"  
5361. 5518  
/note="L1MA4 repeat: matches 6152. .6300 of consensus"  
5508. 6164  
/note="L1MD1 repeat: matches 5423. .6091 of consensus"  
6176. 6245  
/note="L1MD1 repeat: matches 5423. .6091 of consensus"  
6247. 6528  
/note="AluSg repeat: matches 15. .291 of consensus"  
6529. 6674  
/note="L2 repeat: matches 15. .291 of consensus"  
6741. 6874  
/note="L1MD1 repeat: matches 6090. .6223 of consensus"  
7225. 7278  
/note="MIR repeat: matches 36. .85 of consensus"  
8137. 8448  
/note="AluS repeat: matches 1. .312 of consensus"  
9092. 9137  
/note="L2 repeat: matches 1. .312 of consensus"  
9229. 9326  
/note="L2 repeat: matches 2574. .2670 of consensus"  
9669. 9732  
/note="MIR repeat: matches 87. .148 of consensus"  
10162. 10365  
/note="MIR repeat: matches 8. .222 of consensus"  
10379. 10505  
/note="L2 repeat: matches 2586. .2710 of consensus"  
10654. 10814  
/note="MER8 repeat: matches 9. .176 of consensus"  
11454. 11752

repeat\_region  
11790. 12171  
/note="AluIb repeat: matches 1. .302 of consensus"  
12172. 12362  
/note="L1M4 repeat: matches 1258. .1640 of consensus"  
12363. 12775  
/note="MER53 repeat: matches 1. .189 of consensus"  
12833. 13130  
/note="L1M4 repeat: matches 1640. .2049 of consensus"  
13476. 14035  
/note="AluI repeat: matches 1. .291 of consensus"  
14036. 14330  
/note="L1M4 repeat: matches 2224. .2742 of consensus"  
14331. 14466  
/note="AluI repeat: matches 1. .301 of consensus"  
14467. 14600  
/note="L1M4 repeat: matches 2742. .2867 of consensus"  
14601. 14885  
/note="FLAMC repeat: matches 1. .133 of consensus"  
14893. 15096  
/note="L1M4 repeat: matches 2867. .3278 of consensus"  
15159. 15484  
/note="MTRD repeat: matches 286. .505 of consensus"  
15487. 16041  
/note="MTRD repeat: matches 6. .323 of consensus"  
16069. 16773  
/note="L1MC4 repeat: matches 7451. .7973 of consensus"  
16774. 17068  
/note="L1ME3A repeat: matches 3211. .3943 of consensus"  
17071. 17384  
/note="AluSg repeat: matches 3. .297 of consensus"  
17385. 17451  
/note="AluSg repeat: matches 1. .310 of consensus"  
17567. 17775  
/note="L1ME3A repeat: matches 3943. .3955 of consensus"  
17912. 18219  
/note="L1ME3A repeat: matches 4012. .4208 of consensus"  
18401. 18922  
/note="AluSg repeat: matches 1. .313 of consensus"  
18923. 19226  
/note="L1ME3A repeat: matches 4315. .4848 of consensus"  
19227. 19530  
/note="AluS repeat: matches 1. .304 of consensus"  
19531. 19897  
/note="L1ME3A repeat: matches 4848. .4586 of consensus"  
19898. 20698  
/note="L1P81 repeat: matches 5781. .6155 of consensus"  
21061. 21150  
/note="L1ME3A repeat: matches 4585. .5403 of consensus"  
21218. 21390  
/note="MIR repeat: matches 115. .213 of consensus"  
21511. 21767  
/note="L2 repeat: matches 1626. .1814 of consensus"  
21771. 21969  
/note="L1MC4 repeat: matches 6664. .6940 of consensus"  
21970. 22008  
/note="L1MC2 repeat: matches 3850. .4054 of consensus"  
22009. 22354  
/note="L2 repeat: matches 1. .39 of consensus"  
22355. 22663  
/note="L1MC2 repeat: matches 4054. .4410 of consensus"  
22664. 23540  
/note="AluS repeat: matches 1. .309 of consensus"  
23552. 23730  
/note="L1MC2 repeat: matches 4410. .5261 of consensus"  
23731. 24657  
/note="FAM repeat: matches -135. .43 of consensus"  
24659. 24758  
/note="L1MC2 repeat: matches 5278. .6319 of consensus"  
24759. 25029  
/note="AluI/FLAM repeat: matches 2. .77 of consensus"  
25872. 26333  
/note="L1MC4 repeat: matches 7461. .7777 of consensus"  
/note="L1MB5 repeat: matches 4928. .5368 of consensus"

repeat\_region 26334..26622  
/note="LIMB8 repeat: matches 5995..6287 of consensus"  
repeat\_region 26623..26774  
/note="LIMB5 repeat: matches 5368..5512 of consensus"  
repeat\_region 26775..27077  
/note="LIMB5 repeat: matches 5368..5512 of consensus"  
repeat\_region 27078..27656  
/note="LIMB5 repeat: matches 5312..6176 of consensus"  
repeat\_region 27713..27944  
/note="LIMB5 repeat: matches 5312..6176 of consensus"  
repeat\_region 27959..28020  
/note="LIMB3 repeat: matches 6941..7003 of consensus"  
repeat\_region 28028..28269  
/note="LIMB3 repeat: matches 6941..7003 of consensus"  
repeat\_region 28146..28221  
/note="LIMB3 repeat: matches 6941..7003 of consensus"  
misc\_feature 28146..28221  
/note="LIMB3 repeat: matches 6941..7003 of consensus"  
repeat\_region 28512..28663  
/note="LIMB3 repeat: matches 6941..7003 of consensus"  
repeat\_region 29643..30184  
/note="LIMB3 repeat: matches 6941..7003 of consensus"  
repeat\_region 30368..30450  
/note="LIMB3 repeat: matches 6941..7003 of consensus"  
repeat\_region 31248..31409  
/note="LIMB3 repeat: matches 6941..7003 of consensus"  
repeat\_region 31416..31494  
/note="LIMB3 repeat: matches 6941..7003 of consensus"  
repeat\_region 31416..31494  
/note="LIMB3 repeat: matches 6941..7003 of consensus"

Query Match 13.0% Score 260; DB 9; Length 129747;

Best Local Similarity 68.4% Pred. No. 3.6e-61; Matches 547; Conservative 0; Mismatches 140; Indels 113; Gaps 9;

1221 TGCAGAGAGCCAGATGAGAGTGGTGGCTTCCCTATTTTGTCCGAGATGCAT 1280  
1221 TGCAGAGAGCCAGATGAGAGTGGTGGCTTCCCTATTTTGTCCGAGATGCAT 1280  
25030 TTCAATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 25088  
1281 CTTAGGAG 1340  
25089 CTTAGGAG 25148  
1341 AAGGAG 1400  
25149 AAGGAG 25204  
1401 CATTGAG 1460  
25205 CATTGAG 25254  
1461 GACCGAG 1520  
25255 GACCGAG 25254  
1521 AGTCTGACCAATGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1580  
25255 AGTCTGACCAATGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25299  
1581 GATTGAG 1637  
25300 GATTGAG 25359  
1638 GATTGAG 1695  
25360 GATTGAG 25419  
1696 CTTAG 1746  
25420 CTTAG 25479  
1747 GCT-CTGTTGAG 1805  
25480 GCT-CTGTTGAG 25539  
1806 CTTGAG 1865

Db 25540 CTTGAG 25599  
Qy 1866 GGCAG 1924  
Db 25600 GGCAG 25659  
Qy 1925 AGCAG 1984  
Db 25660 AGCAG 25712  
Qy 1985 AAAAAAAAAAAAAAAAAA 2004  
Db 25713 TAAACTTTTAAAGAA 25732

RESULT 10  
G23637/c 391 bp DNA linear SRS 31-MAY-1996  
LOCUS G23637 human SRS WI-13434, sequence tagged site.  
DEFINITION G23637  
ACCESSION G23637  
VERSION G23637.1 GI:1343963  
KEYWORDS SRS; SRS sequence; primer; sequence tagged site.  
SOURCE human SRS derived from sequences in dbEST and the Unigene collection.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 391)  
AUTHORS Hudson, T.  
TITLE Whitehead Institute/MIT Center for Genome Research; Physically Mapped STS  
JOURNAL Unpublished

CONTACT: Thomas Hudson  
Whitehead Institute/MIT Center for Genome Research  
Whitehead Institute for Biomedical Research  
9 Cambridge Center, Cambridge MA 02142 USA  
Tel: 617 252 1900  
Fax: 617 252 1902  
Email: thudson@genome.wi.mit.edu  
PRIMER A: TACATGAATATCTTCATCTCTC  
PRIMER B: CTGTGACAGAGTGCTTGA  
STS size: 150  
PCR Profile:  
Presoak:  
Denaturation:  
Annealing: 56 degrees C  
Polymerization:  
PCR Cycles: 35  
Thermal Cycler:  
Protocol:  
Template: 10 ng  
Primer: each 5 pm  
dNTPs: each 4 mM  
Tag Polymerase: 0.025 units/ul  
Total Vol: 20 ul  
Buffer:  
MgCl2: 1.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 9.3  
Derived from dbEST (genbank accession R40580).  
Location/Qualifiers  
1..391  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="198.8 cr from top of Chr3 linkage group"  
STS  
primer\_bind 32..181  
primer\_bind complement(162..181)

BASE COUNT 87 a 93 c 106 g 99 t 6 others  
ORIGIN

Query Match 10.6%; Score 212.8; DB 11; Length 391;  
Best Local Similarity 92.6%; Pred. No. 2.8e-48;  
Matches 289; Conservative 0; Mismatches 14; Indels 9; Gaps 6;

OY 1704 GGGAGGTGACATGACC-AGCCAGATGATCCTGACAGAGTCTG---CTGTGTTGCCA 1758  
|||||  
DB 321 GGAAGCTGACATGACCCAGCCAGCCAGTGCACAGAGTCTGACAGTCTGTTGCC 262  
OY 1759 GGGCTTGTATTATAG-CCATGATCAGATGTGTGACAGTCTTCTGAG-CTTGAGACCA 1816  
|||||  
DB 261 GGCTCCTGTTATAGCCATGATCAGATGTGTGACAGTCTTCTGAGCCCTGAGACCA 202  
OY 1817 CGGTCACTTGTACGTCTCTGTGACACAGAGTCTTGAGGCATCTCAGCAGCCTCAG 1876  
|||||  
DB 201 CGGTCACTTGTACGTCTCTGTGACACAGAGTCTTGAGGCATCTCAGCAGCCTCAG 142  
OY 1877 CCCAGCTTACCTGCTGCTTGTGACTTGCTT-CTAGCATACCTGGGCAAGCAGGGTGGG 1935  
|||||  
DB 141 CCCAGCTTACCTGCTGCTTGTGACTTGCTTCTTCTAGCATACCTGGGCAAGCAGGGTGGG 82  
OY 1936 GAATGAGAGATAGACATGGGATGTATGAGAGAGATGAGAGATTTTCCGAAAAA 1995  
|||||  
DB 81 GAATGAGAGATAG-CATGGATGTATGAGAGAGATGAGAGATTTTCAATAATAAAT 23  
OY 1996 AAAAAAAAAA 2007  
|||||  
DB 22 TAAAAAAAAA 11

## RESULT 11

AX350344 AX350344 7542 bp DNA linear PAT 06-FEB-2002  
LOCUS Sequence 3 from Patent WO0192492.  
DEFINITION AX350344  
ACCESSION AX350344  
VERSION AX350344.1 GI:18616006  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (sites)  
AUTHORS Yan, C.C., Wei, M.H., Ketchum, K.C., Merkulov, G.C. and Beasley, E.M.  
TITLE Isolated human kinase proteins, nucleic acid molecules encoding  
JOURNAL human kinase proteins, and uses thereof  
Patent: WO 0192492-A 3 06-DEC-2001;

FEATURES  
source 1..7542  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1612 a 1977 c 2156 g 1797 t  
ORIGIN

Query Match 8.8%; Score 176.4; DB 6; Length 7542;  
Best Local Similarity 95.1%; Pred. No. 6.5e-38;  
Matches 194; Conservative 0; Mismatches 6; Indels 4; Gaps 1;

OY 1 GTGCGTGAGCGAATTGTAACAGGAGGAGGCGGCGG---CTCTACGAGCGCGGAC 56  
|||||  
DB 7302 GTGCTCGGAGCCAAATTTGAACAGGAGGAGGCGGCGGCGCTCTACGAGCGCGAC 7361  
OY 57 CTGTACAGATTTCTTGGCTGCTGGGCCCTTGAGTCCAGCATGCGCTTCCGTGC 116  
|||||  
DB 7362 CTGTACAGATTTCTTGGCTGCTGGGCCCTTGAGTCCAGCATGCGCTTCCGTGC 7421  
OY 117 TCTGTCACTATCTGCTCGACTTCTTGATCACTCCCGGAGCGTGGCGGCATCCACTG 176  
|||||  
DB 7422 TCTGTCACTATCTGCTCGACTTCTTGATCACTCCCGGAGCGTGGCGGCATCCACTG 7481

OY 177 CGGCACACCTTCCACTTGCACTG 200  
DB 7482 CGGCACACCTTCCACTTGCACTG 7505  
|||||

## RESULT 12

AC099330 163424 bp DNA linear HTG 09-NOV-2001  
LOCUS Homo sapiens chromosome 3 clone RP11-78010, WORKING DRAFT SEQUENCE,  
DEFINITION 8 unordered pieces.  
AC099330 AC068701  
AC099330.1 GI:16874864  
VERSION  
KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 163424)  
AUTHORS Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 163424)  
AUTHORS Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
TITLE Direct Submission  
JOURNAL Submitted (09-NOV-2001) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
On Nov 9, 2001 this sequence version replaced gi:8247914.

## COMMENT

Genome Center  
Center: University of Washington Genome Center  
Center Code: UMG  
Web site: <http://www.genome.washington.edu>  
Contact: [umgchigs@u.washington.edu](mailto:umgchigs@u.washington.edu)  
Drafting Center: BCM  
----- Project Information -----  
Center project name: chr-3  
Center clone name: RP11-78010 (bc0618)  
----- Summary Statistics -----  
Sequencing vector: unknown; 42% of reads  
Chemistry: Dye-terminator Big Dye; 92% of reads  
Chemistry: Dye-terminator Big Dye; 8% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 158969 bases at least Q40  
Consensus quality: 160979 bases at least Q40  
Consensus quality: 162023 bases at least Q20  
Insert size: 202800; 21.9% error; agarose-IP  
Insert size: 162724; sum-of-contrigs  
Quality coverage: 9.0x in Q20 bases; agarose-IP  
Quality coverage: 11.2x in Q20 bases; sum-of-contrigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 8 contrigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contrigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 4608: contrig of 4608 bp in length  
\* 4609 4708: gap of unknown length  
\* 4709 10289: contrig of 5581 bp in length  
\* 10290 10389: gap of unknown length  
\* 10390 17105: contrig of 6716 bp in length  
\* 17106 17205: gap of unknown length  
\* 17206 45100: contrig of 27895 bp in length  
\* 45101 45200: gap of unknown length  
\* 45201 65174: contrig of 19974 bp in length  
\* 65175 65274: gap of unknown length  
\* 65275 90719: contrig of 23445 bp in length  
\* 90720 90820: gap of unknown length  
\* 90820 119377: contrig of 28558 bp in length  
\* 119378 119477: gap of unknown length



```

FEATURES          * 119478 163424: contig of 43947 bp in length.
source             1. 163424
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="3"
                  /clone="RP11-78010"
                  /clone_11b="RPci human BAC library 11"
misc_feature       1. 4608
                  /note="assembly_name:Contig74"
misc_feature       4709..10289
                  /note="assembly_name:Contig75"
misc_feature       10390..17105
                  /note="assembly_name:Contig76"
misc_feature       17206..45100
                  /note="assembly_name:Contig77"
misc_feature       45201..65174
                  /note="assembly_name:Contig78"
misc_feature       65275..90719
                  /note="assembly_name:Contig79"
misc_feature       90820..119377
                  /note="assembly_name:Contig80"
misc_feature       119478..163424
                  /note="assembly_name:Contig81"
BASE COUNT      40950 a 40937 c 39779 g 40912 t      846 others
ORIGIN
Query Match       8.8%; Score 176.4; DB 2; Length 163424;
Best Local Similarity 95.1%; Pred. No. 9.9e-38;
Matches 194; Conservative 0; Mismatches 6; Indels 4; Gaps 1;

QY 1 GTGCGGTGAGCGCAATTGTAAGACAGCGGAGCGGGGCGC-----CTCTAGCAGCGCGGAC 56
Db 149275 GTGCTCGGAGCCAAATTTGAACAGAGCGGCGGCGCGCTCTAGCAGCGGAC 149334

QY 57 CTGTACAGATTCTTTGGCTGCTGCGGCGCCCTTGAGTCCAGCCATCATGCTATCCGTGC 116
Db 149335 CTGTACAGATTCTTTGGCTGCTGCGGCGCCCTTGAGTCCAGCCATCATGCTATCCGTGC 149394

QY 117 TCTGTGCATATCTGCTCCGACTCTTTCATCTCCCGGAGCGTGGCCGCATCCACTG 176
Db 149395 TCTGTGCATATCTGCTCCGACTCTTTCATCTCCCGGAGCGTGGCCGCATCCACTG 149454

QY 177 CGGCGACACCTTCACACTTGCAGTG 200
Db 149455 CGGCGACACCTTCACACTTGCAGTG 149478

RESULT 13
AL353694_2/c
WPCOMMENT
Sequence split into 4 fragments LOCUS AL353694 Accession AL353694
Fragment Name      Begin      End
AL353694_0         1      110000
AL353694_1        100001    210000
AL353694_2        200001    310000
AL353694_3        300001    356330
Continuation (3 of 4) of AL353694 from base 200001 (AL353694 Homo sapiens chromosome 6 c

Query Match       6.5%; Score 130; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 9.4e-25;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 501 CACACTGAAA 510
Db 62626 CACACTGAAA 62617

RESULT 14
AC099330/163424 bp DNA linear HTG 09-NOV-2001
LOCUS Homo sapiens chromosome 3 clone RP11-78010, WORKING DRAFT SEQUENCE,
DEFINITION 8 unordered pieces.
ACCESSION AC099330.1 GI:16874864
VERSION AC099330.1 GI:16874864
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULFILLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 163424)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submission
Unpublished
2 (bases 1 to 163424)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Submitted (09-NOV-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Nov 9, 2001 this sequence version replaced gi:8247914.

***** Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgs@u.washington.edu
Drafting Center: BCM

***** Project Information
Center project name: chr-3
Center clone name: RP11-78010 (bc0618)

***** Summary Statistics
Sequencing vector: unknown; 42% of reads
Chemistry: Dye-terminator ET; 92% of reads
Chemistry: Dye-terminator Big Dye; 8% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 158969 bases at least Q40
Consensus quality: 160979 bases at least Q30
Consensus quality: 162023 bases at least Q20
Insert size: 202800; 21.9% error; agarose-fp
Insert size: 162724; sum-of-ctrligs
Quality coverage: 9.0x in Q20 bases; agarose-fp
Quality coverage: 11.2x in Q20 bases; sum-of-ctrligs

```

```

***** NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
4608: contig of 4608 bp in length
4708: gap of unknown length
4709 10289: contig of 5581 bp in length
10290 10389: gap of unknown length
10390 17105: contig of 6716 bp in length
17106 17205: gap of unknown length
17206 45100: contig of 27895 bp in length
45101 45201: gap of unknown length
45202 65174: contig of 19974 bp in length
65175 65274: gap of unknown length
65275 90719: contig of 25445 bp in length
90720 90819: gap of unknown length
90820 119377: contig of 28558 bp in length
119378 119478: gap of unknown length
119479 163424: contig of 43947 bp in length.

```

FEATURES	SOURCE	Location/Qualifiers
misc_feature		1..163424
misc_feature		/organism="Homo sapiens"
misc_feature		/db_xref="taxon:9606"
misc_feature		/chromosome="3"
misc_feature		/clone="RP11-78010"
misc_feature		/clone_lib="RPCI human BAC library 11"
misc_feature		1..4608
misc_feature		/note="assembly_name:Contig74"
misc_feature		4709..10289
misc_feature		/note="assembly_name:Contig75"
misc_feature		10390..117105
misc_feature		/note="assembly_name:Contig76"
misc_feature		17206..45100
misc_feature		/note="assembly_name:Contig77"
misc_feature		45201..65174
misc_feature		/note="assembly_name:Contig78"
misc_feature		65275..90719
misc_feature		/note="assembly_name:Contig79"
misc_feature		90820..119377
misc_feature		/note="assembly_name:Contig80"
misc_feature		119478..163424
misc_feature		/note="assembly_name:Contig81"
BASE COUNT	40950 a 40937 c 39779 g 40912 t	846 others
ORIGIN		
Query Match	6.5%; Score 130; DB 2; Length 163424;	
Best Local Similarity	100.0%; Pred. No. 1e-24;	
Matches 130; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
OY 361	AGACAGAGGAAGACAGACAGCCAGGTCATCTGCACACTCTGGCGGATAGCCTGGAGACA	440
Db 57007	AGACAGAGGAAGAACAGACAGCCAGGTCATCTGCACACTCTGGCGGATAGCCTGGAGACA	56948
OY 441	ACGCAATGCTACTGTGATCTCTGACAGAGCCCTGGGCAAGGCCAGATGCTGTGCTC	500
Db 56947	ACGCAATGCTACTGTGATCTCTGACAGAGCCCTGGGCAAGGCCAGATGCTGTGCTC	56888
OY 501	CACACTGAAA 510	
Db 56887	CACACTGAAA 56878	
RESULT 15		
FR165K09/c		
LOCUS	FR165K09	39410 bp DNA linear VRT 23-JAN-2002
DEFINITION	Fugu rubripes cosmid 165K09 DNA for GRM7, TRIP, Sand, PRGR3 genes	
ACCESSION	AF010317	
VERSION	AF010317.1 GI:3928163	
KEYWORDS	GRM7 gene; metabotropic glutamate receptor 7; plasminogen related growth factor receptor 3; PRGR3 gene; Sand gene; TRAF Interacting protein; trip gene.	
SOURCE	Takifugu rubripes.	
ORGANISM	Takifugu rubripes.	
REFERENCE	1 (sites)	
AUTHORS	Cottage,A.J., Clark,M., Hawker,K., Umranta,Y., Wheller,D., Bishop,M. and Elgar,G.	
TITLE	Three receptor genes for plasminogen related growth factors in the genome of the puffer fish Fugu rubripes	
JOURNAL	FEBS Lett. 443 (3), 370-374 (1995)	
MEDLINE	99148833	
PUBMED	10025966	
REFERENCE	2 (sites)	
AUTHORS	Cottage,A.J., Edwards,Y.J. and Elgar,G.	
TITLE	SAND, a new protein family: from nucleic acid to protein structure and function prediction	
JOURNAL	Compar. Funct. Genom. 2, 226-235 (2001)	
REFERENCE	3 (bases 1 to 39410)	

FEATURES	source	location/Qualifiers
AUTHORS	Cottage A.J.	Direct Submission
TITLE	Submitted (15-AUG-1998)	Cottage A.J., HGMPC, Computing, Hinxton,
JOURNAL	Cambridge, CB10 1SB, UK	
FEATURES		location/Qualifiers
source	1..39410	/organism="Takifugu rubripes"
		/db_xref="taxon:31033"
		/clone="cosmid:165K09"
mrna	join(<261..401,496..696,1891..2030,2155..2778,2854..3165,4227..>4466)	/gene="GRM7"
gene	261..4466	/gene="GRM7"
	join(<261..401,496..696,1891..2030,2155..2778,2854..3165,4227..>4466)	/gene="GRM7"
	/product="metabotropic glutamate receptor 7"	
	/protein_id="CA009083.1"	
	/db_xref="GI:3928164"	
	/db_xref="SPTREMBL:O9PW00"	
	/translation="PDQYFTSRSLENNRRNINWFAEFWEDDFKCKLTRPKIKELGRKCRDEEDIRSDSOYEOEGKVOPIIDAVYAAHAALSHMHDLCGSMGVCKMPVGGRLLOYTHGVNFGNSAGTGVWFMENNGDAPGRYDIFQYQMSNISNPGYRNICQWNLRLNLEPMRNSGGDRKIPESVCSPPCESGSRKKMKVGVPCMCHELDGTYAIRNDPIVRYMQRPLKNTGCRPTPIIKLEWSSPMALIPVFLALIGLATGVATPIRNDPIVRSAGRELSYLLGLFIPLIPLFLMAIEPSVAVCARFLFLGMCISASMLTQNR IYVIFDQKSVTPPKPISPTSQLITTFILISVOLGLVIFMGVMPPHIITDEQKP PNEPFRGVFKCDMSDLSTILICSVLIMITCTVYASISGVPTEFNFKRIGTFMY TNCIVLAPVPIPIFGTAOSTEKKFITOTTLTVLSMSLSATVSGLMLIIPVYIITFHE QNVOKKRSFKAVAVATVSTRLSKSDKONGESKIEPDRSQ"	
exon	<261..401	/gene="GRM7"
	/number=1	
intron	402..495	/gene="GRM7"
	/number=1	
exon	496..696	/gene="GRM7"
	/number=2	
intron	697..1890	/gene="GRM7"
	/number=2	
exon	1891..2030	/gene="GRM7"
	/number=3	
intron	2031..2154	/gene="GRM7"
	/number=3	
exon	2155..2778	/gene="GRM7"
	/number=4	
intron	2779..2853	/gene="GRM7"
	/number=4	
exon	2854..3165	/gene="GRM7"
	/number=5	
intron	3166..4226	/gene="GRM7"
	/number=5	
exon	4227..>4466	/gene="GRM7"
	/number=6	
mrna	complement(join(<4660..4801,6176..6226,6468..6521,6616..6743,6799..6931,7017..7106,7205..7292,7382..7495,7643..7737,7816..7943,8029..8068,8194..8277,8358..8415,8496..>8593))	/gene="GRM7"
	/gene="TRIP"	
gene	complement(4660..8593)	/gene="TRIP"
	/gene="TRIP"	
CDS	complement(join(4660..4801,6176..6226,6468..6521,	



```
QY 640 GTGAGAGATGATCCGACATGGTGTGGACAGTACCGGTGAACAGCTGTGTG 639
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1345 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1286
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 700 TACTGTCTCTCTCAAGAAAGTACGACATCTAAAGAGCAGCAGCCTCAGGG 759
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1285 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1226
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 760 GAGGTGCTCAGCAGCAGGAGGATTTGTTTCCCTCCAGAGCAGTTCGACAGTC 819
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1225 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1166
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 820 TACTCTGAATTCGATCAGCCAGTTCGATCAGTCCAGCAGCAGTTCACAGAT 879
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1165 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1106
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 880 GCTGACAGAGAATCATGACCTGAAAAAGACGTAACGAT 920
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1105 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1065
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 17
AC067854/C 203984 bp DNA linear HTG 04-JAN-2002
LOCUS Homo sapiens chromosome 8 clone RP11-513H8 map 8, ** SEQUENCING IN
DEFINITION PROGRESS ***, 1 ordered pieces.
AC067854
AC067854.13 GI:18056699
KEYWORDS HTG; HTGS_PHASE2; HTGS_FUOTOP; HTGS_ACTIVEPTIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 203984)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-513H8
Unpublished
2 (bases 1 to 203984)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boucknight,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galsagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hages,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPeeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,J., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pletre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testage,S., Theodore,J., Tirrell,A., Travers,M., Tigilillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
```

COMMENT  
TITLE  
JOURNAL  
Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 4, 2002 this sequence version replaced g1:18034761.  
All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green,P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information

```
Center project name: L9308
Center clone name: 513_H_8
*****
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 203984: contig of 203984 bp in length.
Location/Qualifiers
1. 203984
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-513H8"
/clone_lib="RP11-11 Human Male BAC"
BASE COUNT 61400 a 39486 c 39964 g 63134 t
ORIGIN
```

Query Match 3.0%; Score 60.6; DB 2; Length 203984;  
Best Local Similarity 44.8%; Pred. No. 2.9e-05;  
Matches 280; Conservative 0; Mismatches 339; Indels 6; Gaps 1;

```
QY 305 AGGAGGAGGAGAAATGCTTGTGATCGAATCTTAAAGATGAACATGACATGACAG 364
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16618 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16559
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 365 CCAGCTTTCCCAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 424
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15558 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16499
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 425 GGGATACGCTGTAAGAAACGCAATGCTACTGTATCTTCGACAGCCTTGAGCAGG 484
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16498 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16439
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 485 CCGAGATCTGTCTCTCCACACTGAAAAAGCAGATGAAGTCTTAGAGCAGCAGCAGT 544
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16438 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16385
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 545 AGACAAACAGCAAGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 604
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16384 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16325
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 605 AGATTGAGCTTCTAATCTCAGAGCAGCTCCCTGAGGTGAGAGATGATCCGAGCATGG 664
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16324 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16265
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 665 GTGTGGACAGTACGCGGTGGAACAGCTGCTGTGATCTGTCTCTCAAGAAAGT 724
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16264 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16205
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 725 ACAGAAATCTTAAAGAGCAGCAGGAGCCTCAGGAGAGTGTGCTGACAACTGAGAGA 784
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16204 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16145
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 785 ATTTGTTTCTCCAGAGCAAGTTCGAGACACTTACTCTGAATTCGATCCAGCAAGT 844
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16144 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16085
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 845 TAGAACTGAAGTACGCCAAGAGCACTTACAGAGTCTGCAAGCAAGAAATCATGAGCTTGA 904
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16084 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16025
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 905 AAAAGAGCTTACGATCTTCGACGA 929
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16024 AGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16000
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 18  
AC073495/c  
LOCUS  
DEFINITION  
AC073495  
AC073495.12 GI:10801940  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AC073495 318930 bp DNA linear HTG 15-OCT-2000  
Mus musculus chromosome 11 clone RP23-345K4, WORKING DRAFT  
SEQUENCE, 64 unordered pieces.  
AC073495  
AC073495.12 GI:10801940  
HTG, HTGS, PHASE1, HTGS\_DRAFT.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 318930)  
Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,  
Deckerich, D., Thomas, S., Okwuonu, G., Carlock, C., Garner, T.,  
Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J.,  
Buhay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,  
Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S.,  
Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Gill, R.,  
Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hoques, M.,  
Kosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,  
Kovari, C., Liu, J., Liu, W., Louised, H., Lozado, R.J., Martin, R.,  
Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S.,  
Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogul, M., Parish, B.,  
Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S.,  
Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A.,  
Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstock, G.,  
Worley, K. and GDBs, R.  
Direct Submission  
Unpublished  
2 (bases 1 to 318930)  
Worley, K.C.  
Direct Submission  
Submitted (19-JUN-2000) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Oct 15, 2000 this sequence version replaced gi:10800205.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: MACH  
Center clone name: RP23-345K4  
----- Summary Statistics  
Sequencing vector: M13; 108821  
Chemistry: Dye-Primer Bodipy: 71% of reads  
Chemistry: Dye-terminator Big Dye: 29% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 282577 bases at least Q40  
Consensus quality: 299603 bases at least Q30  
Consensus quality: 309271 bases at least Q20  
Estimated insert size: 298164; sum-of-coverage estimation  
Quality coverage: 0x in Q20 bases; agarose-gel estimation  
Quality coverage: 5.1x in Q20 bases; sum-of-coverage estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_dirft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_dirft_data.html))  
\* NOTE: This sequence may represent more than one clone.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 64 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 100449: contig of 100449 bp in length  
\* 100450 100549: gap of unknown length  
\* 100550 143847: contig of 43298 bp in length  
\* 143848 143947: gap of unknown length

143948 152393: contig of 8446 bp in length  
\* 152394 152493: gap of unknown length  
\* 152494 160427: contig of 7933 bp in length  
\* 160427 160526: gap of unknown length  
\* 160527 166643: contig of 6117 bp in length  
\* 166644 166744: gap of unknown length  
\* 166744 173484: contig of 6741 bp in length  
\* 173485 173585: gap of unknown length  
\* 173585 177408: contig of 3824 bp in length  
\* 177409 177508: gap of unknown length  
\* 177509 183311: contig of 5802 bp in length  
\* 183311 188977: contig of 5567 bp in length  
\* 188978 189078: gap of unknown length  
\* 189078 196629: contig of 7552 bp in length  
\* 196630 196730: gap of unknown length  
\* 196730 202134: contig of 5405 bp in length  
\* 202135 202235: gap of unknown length  
\* 202235 206964: contig of 4729 bp in length  
\* 206964 207063: gap of unknown length  
\* 207064 212314: contig of 5251 bp in length  
\* 212315 212414: gap of unknown length  
\* 212415 216415: contig of 4001 bp in length  
\* 216416 216515: gap of unknown length  
\* 216516 219625: contig of 3110 bp in length  
\* 219626 219725: gap of unknown length  
\* 219726 224004: contig of 4279 bp in length  
\* 224005 224104: gap of unknown length  
\* 224105 227363: contig of 3259 bp in length  
\* 227364 227463: gap of unknown length  
\* 227464 232010: contig of 4547 bp in length  
\* 232011 232110: gap of unknown length  
\* 232111 234955: contig of 2845 bp in length  
\* 234956 235055: gap of unknown length  
\* 235056 238628: contig of 3573 bp in length  
\* 238629 242373: gap of unknown length  
\* 242374 242473: contig of 3445 bp in length  
\* 242474 244474: gap of unknown length  
\* 244475 244574: contig of 2001 bp in length  
\* 244575 247650: contig of 3076 bp in length  
\* 247651 247750: gap of unknown length  
\* 247751 250706: contig of 2956 bp in length  
\* 250707 250806: gap of unknown length  
\* 250807 253646: contig of 2840 bp in length  
\* 253647 253746: gap of unknown length  
\* 253747 255897: contig of 2151 bp in length  
\* 255898 255997: gap of unknown length  
\* 255998 258504: contig of 2507 bp in length  
\* 258505 258604: gap of unknown length  
\* 258605 260676: contig of 2072 bp in length  
\* 260677 260776: gap of unknown length  
\* 260777 263190: contig of 2414 bp in length  
\* 263191 263290: gap of unknown length  
\* 263291 264984: contig of 1694 bp in length  
\* 264985 265084: gap of unknown length  
\* 265085 266722: contig of 1638 bp in length  
\* 266723 266822: gap of unknown length  
\* 266823 268472: contig of 1650 bp in length  
\* 268473 268572: gap of unknown length  
\* 268573 270376: contig of 1804 bp in length  
\* 270377 270476: gap of unknown length  
\* 270477 272058: contig of 1582 bp in length  
\* 272059 272158: gap of unknown length  
\* 272159 274342: contig of 2084 bp in length  
\* 274343 274442: gap of unknown length  
\* 274443 276862: contig of 2520 bp in length  
\* 276863 276962: gap of unknown length  
\* 276963 279114: contig of 2152 bp in length  
\* 279115 279214: gap of unknown length  
\* 279215 280875: contig of 1661 bp in length  
\* 280876 280975: gap of unknown length  
\* 280976 282744: contig of 1769 bp in length

28275	282844:	gap of unknown length
282845	2828991:	contlg of 1147 bp in length
282932	284091:	gap of unknown length
284092	285944:	contlg of 1856 bp in length
2840948	286044:	gap of unknown length
286048	287985:	contlg of 1938 bp in length
287986	288085:	gap of unknown length
288086	288998:	contlg of 1813 bp in length
288989	289998:	gap of unknown length
289999	291415:	contlg of 1417 bp in length
291416	291515:	gap of unknown length
291516	292676:	contlg of 1161 bp in length
292777	292776:	gap of unknown length
292777	294223:	contlg of 1447 bp in length
294224	294323:	gap of unknown length
294324	295422:	contlg of 1039 bp in length
295423	295522:	gap of unknown length
295523	296939:	contlg of 1417 bp in length
296940	297039:	gap of unknown length
297040	298270:	contlg of 1231 bp in length
298271	298370:	gap of unknown length
298371	299936:	contlg of 1566 bp in length
299937	300036:	gap of unknown length
300037	301414:	contlg of 1378 bp in length
301415	301514:	gap of unknown length
301515	303105:	contlg of 1591 bp in length
303106	303965:	gap of unknown length
303966	304965:	contlg of 1760 bp in length
304966	305065:	gap of unknown length
305066	306362:	contlg of 1297 bp in length
306363	306462:	gap of unknown length
306463	307463:	contlg of 1001 bp in length
307464	307563:	gap of unknown length
307564	308714:	contlg of 1151 bp in length
308715	308814:	gap of unknown length
308815	309956:	contlg of 1142 bp in length
309957	310056:	gap of unknown length
310057	311344:	contlg of 1288 bp in length
311345	311444:	gap of unknown length
311445	312654:	contlg of 1210 bp in length
312655	312754:	gap of unknown length
312755	313811:	contlg of 1057 bp in length
313812	313911:	gap of unknown length
313912	315044:	contlg of 1133 bp in length
315045	315144:	gap of unknown length
315145	316580:	contlg of 1436 bp in length
316581	316680:	gap of unknown length
316681	317778:	contlg of 1098 bp in length
317779	318780:	gap of unknown length
318789	318930:	contlg of 1052 bp in length.

BASE COUNT	84543 a	73358 c	71912 g	82767 t	6350 others
------------	---------	---------	---------	---------	-------------

Query Match	3.0%;	Score 59.6;	DB 2;	Length 318930;
Best Local Similarity	45.2%;	Pred. No. 5.9e-05;		
Matches 218;	Conservative 0;	Mismatches 264;	Indels 0;	Gaps 0;

OY	304	CAGGAGGACGAGAAATGCTTGGATCGAAGAATTCTTAAAGAAATGTGAACATGCATGTCAGA	363
Db	166402	CAGGAGGACGAGAGAGAGACAGAGAGAGAGAGCAGCAGGAGAGAGAGAGAGAGAGAG	166343
OY	364	GCCACGCTTTCCCAGAAAAGCAAGAGAAAAAGACAGACGCCAGTTCATTCGACTCTG	423
Db	166342	GAGGAGGAGAAAGAGAGACAGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	166283
OY	424	CGGGATACGCTGTAAGAACCCAAATGCTACTCTGTGGTATCTCTGCAGCACAGGCTTTGGGCAAG	483
Db	166282	GAGGAGGAGAGAGAGAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA	166223

[illegible]



FEATURES	source
* 41020	41343: gap of unknown length
* 41344	59365: contig of 18022 bp in length
* 59366	59683: gap of unknown length
* 59690	73640: contig of 13951 bp in length
* 73641	73964: gap of unknown length
* 73965	87376: contig of 13412 bp in length
* 87377	87700: gap of unknown length
* 87701	101024: contig of 13324 bp in length
* 101025	101348: gap of unknown length
* 101349	112980: contig of 11632 bp in length
* 112981	113304: gap of unknown length
* 113305	124819: contig of 11515 bp in length
* 124820	125143: gap of unknown length
* 125144	136091: contig of 10948 bp in length
* 136092	136415: gap of unknown length
* 136416	144125: contig of 7710 bp in length
* 144126	144449: gap of unknown length
* 144450	152005: contig of 7556 bp in length
* 152006	152329: gap of unknown length
* 152330	159587: contig of 7258 bp in length
* 159588	159911: gap of unknown length
* 159912	167134: contig of 7223 bp in length
* 167135	167458: gap of unknown length
* 167459	173757: contig of 6299 bp in length
* 173758	174081: gap of unknown length
* 174082	179850: contig of 5769 bp in length
* 179851	180174: gap of unknown length
* 180175	185555: contig of 5381 bp in length
* 185556	185879: gap of unknown length
* 185880	191144: contig of 5265 bp in length
* 191145	191467: gap of unknown length
* 191468	196412: contig of 4945 bp in length
* 196413	196735: gap of unknown length
* 196736	201258: contig of 4523 bp in length
* 201259	201581: gap of unknown length
* 201582	205749: contig of 4168 bp in length
* 205750	206072: gap of unknown length
* 206073	209959: contig of 3887 bp in length
* 209960	210282: gap of unknown length
* 210283	213540: contig of 3258 bp in length
* 213541	213863: gap of unknown length
* 213864	216191: contig of 2328 bp in length.
FEATURES	
source	Location/Qualifiers
	1. .216191
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/clone="RP23-60B7"
BASE COUNT	60680 a 45448 c 44115 g 58724 t 7224 others
ORIGIN	

	Query Match	Similarity	2.9%	Score	59	DB 2	Length	216191
	Best Local	Similarity	43.8%	Pred. No.	8.2e-05			
	Matches	257	Conservative	0	Mismatches	330	Indels	0
								Gaps
								0
0Y	304	CAGGAGGAGGAAATGCTTGGATCGAGATTTCTTAAAGTAATCTGACATGTCAGA	363					
Db	109137	CAGGAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	109078					
0Y	364	GCCAGACCTTTCCAGAAAGACAAAGAGAAAGACAGACAGCAGTCATCTGACACTCTG	423					
Db	109077	CAGCAGGAGCAGAGAGAGAGAGCAGCAGGAGAAAGACAGACAGCAGCAGCAGAAAGAGCAG	109018					
0Y	424	CGGGATAGCCTGGAGGAACGCCAATGCTACTGTGGTATCTCTGACAGAGCCTTTGGCCAG	483					
Db	109017	GAGCAGAGACTGGAGCAGGAGTGAAGGATGAGGAGAGAAAGCAGAGCAGAGAGAGGAGCGAG	108958					
0Y	484	GCCGAGATGCTTGTCTCCACTGTAAAAAGCAGATGAATTAATCTACTTAGCAGCAGCAGAT	543					
Db	108957	GAGCTGGACAGCAGAGAGAGCAGCAGAAACAAAGAGAAAGACAGCAGCTGGAGCAGAGAGAA	108958					
0Y	544	GAGACCAACAAAGCAGCAGAGAGCGCGGCGGCTCAGAGCAGATGAAGACCATGAGC	603					

DB	108897	GAGCAGGAGACTGGACACGAGGACGAGGAGGAGGAGCGAGGAGCAGGAGGAGGAGGAAGAAGCAG	108838
OY	604	CAGATTGAGCTTTCTA <sup>1</sup> CTCCAGACACCAGCTCCCCTGAGGTGGAGGAGATATATCCGAGACTTG	663
DB	108837	GAGCAGGAGGAGAACAGCAGGACGACGAGGAGAAAAAGCAGGAGCAGCAGGAGAGAAACACGAGGACGAG	108778
OY	664	GGTGTGGACACAGTCAGCCGTGGAMACAGCTGGCTGTGTACTGTGTCTCTCAAGAAAGAG	723
DB	108777	GAGGAAGAGCAGGAGACTGGAGCAGGAGCAGGAGGAGGAGCAGGAGGAGGAGGAGGAGGAGCAG	108718
OY	724	TACGGAATCTTAAAGAGGAGCACGGAAGGCCTCAGGGGAGGTGGCTGACAGCTGAGGSAAG	783
DB	108717	GAGCAGGAGGAGAACACAGCAGAGCAGGAGGAAGCAAGCAGGAGCAGGAGGAAGCAGGAGCTG	108658
OY	784	GATTTGTTTTCCCTCAGAACCAACTTGCAGACACTCTACTCTGAATTGATTCAGGCCAAG	843
DB	108657	GAGCTGAAGCAGCAGAGGAGAAACGACGAGCAGGAGGAGGAGCAGGAGCAGACTGTGAGCTGAAGCAG	108598
OY	844	TTAGAGCTGAAGTCAAGCCAGCAAGGAGCTTTACAGATAGCTGCTGACAAGA	890
DB	108597	GAGGAGAACACAGCAGGAGCAGGAGCAGGAGCTGCAGCCAGGAGGAGGAGGAGGA	108551

RESULT	20			
LOCUS	AF410847/C			
DEFINITION	Ovine herpesvirus 2 latency associated antigen (orf73) gene,	2065 bp	DNA	linear
ACCESSION	AF410847			VRL 20-SEP-2001
VERSION	AF410847			
KEYWORDS	AF410847.1	GI:15705839		
SOURCE				
ORGANISM	Ovine herpesvirus 2.			
REFERENCE	Ylvises, dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae.			
AUTHORS	1 (bases 1 to 2065)			
TITLE	Coulter, L.J. and Reid, H.W.			
JOURNAL	Isolation and expression of three open reading frames (ORFs) from			
REFERENCE	ovine herpesvirus 2			
AUTHORS	unpublished			
TITLE	2 (bases 1 to 2065)			
JOURNAL	Coulter, L.J. and Reid, H.W.			
REFERENCE	Direct Submission			
AUTHORS	Submitted (15-AUG-2001)			
TITLE	Virology, Moredun Research Institute,			
JOURNAL	Submitted (15-AUG-2001)			
REFERENCE	Penicuik, Midlothian EH26 0PZ,			
AUTHORS	UK			

[illegible]



QY	724	TACGAGAACTCTAAAGAGCAGCAGAGCCCTCAGGGAGGTGGCTGACACAGTGAGAG	783		Schlossgarten 4, Institut für Klinische und Molekulare, Erlangen 91054, Germany
Db	155474	CACGAGAGAGCAGCAGAGCAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAG	155533	REMARK	Sequence update by submitter
QY	784	GATTGTG 789		COMMENT	On Jul 9, 2001 this sequence version replaced gi:2246466.
Db	155534	AAGTAG 155539		FEATURES	Location/Qualifiers
RESULT	22			source	1..133661
LOCUS	U93872/c				/organism="Human herpesvirus 8"
DEFINITION	U93872	133661 bp	DNA		/db_xref="taxon:37296"
	U93872	Kaposi's sarcoma-associated herpesvirus glycoprotein M, DNA			/note="sequence derived from overlapping Lambda phage clones isolated from two libraries of Sau3a partially digested DNA from Kaposi's sarcoma biopsies"
	U93872	inhibitory protein and v-cyclin genes, complete cds, and tegment			105..959
	U93872	protein gene, partial cds.			/note="ORF K1"
ACCESSION	U93872.2	GI:14627174			/protein_start=1
VERSION	U93872.2	GI:14627174			/protein_id="AAB62619.1"
KEYWORDS					/db_xref="GI:2246494"
SOURCE					/translation="MFLYVCSLAVCFRLSLSLSSDSSPNLCRGVISTPYTITCSBNT
ORGANISM					SLEPTSWYCNDRILRLTQQTFTVDILICNFCSCVGSGHSHSLMTWYPPQVLTQCGQ
					PSNTVYCGGHVTLVLCSTSGNNVYVHMLPGNREYVSQTYNEFTLMNTEGCGYACSG
					LSRSLSNRLCFSAKCANLPEETHVSVSSTGFRFSTNRGRGTTLDVLVMAKSAKSTNL
					HIOVFLVFMVTLVAILGTCGILGTIIIFAHCKQKQKRSNKTVPQQLDYVSLDLCTED
					YVQPVDMV"
					1127..2779
					/note="ORF 04, complement control protein homolog"
					/codon_start=1
					/protein_id="AAB62602.1"
					/db_xref="GI:2246477"
					/translation="MAFLRQTLMTLWTFPMVIGDNEKCSOKTLIGYRLKMSRDGIA
					VERTVFLRCRSGTYYARNTATACLOGGTRSEPTAKNCKSCNPGEIQRNKGVIPIRG
					QDALRTGANTISVCNKGFLVGREYRYCMIGASGMAWSSSPCEKCRKPKTEN
					GGEKPKDYKYENDAVHPECNEGYTLVGHSLACAVANNWTSMPPLCELAGKFPST
					HGYIDPFSLTYKHQSVTFACNDGFLVLSGSPITCNVEMWPLPKCYLIDIDPNN
					SNGRDLHPNENKNGNFORSNVTEPPKPEDTHAATCDTEBPPLILPSEGFN
					ETTSNTITIKOLEDEKTSOPNTHISALTSMAKAGNPNKNTNSDHLIASPTSD
					DAPPSIVSQTPVNYNMAPTRTLTSLIEEGPSNYSSTSEKATASTSHSHKNDGCI
					YTLNKTQDPSNKNPTNSAKSSTKPRVETHNKTTNSPAISITSDSADVPQRPREPTL
					PLFPPPAKNRRLERKOLVIGLTAVALTCGLTTLFHYLFFR"
					3194..6592
					/note="ORF 06, major ssDNA binding protein homolog"
					/codon_start=1
					/protein_id="AAB62603.1"
					/db_xref="GI:2246478"
					/translation="MAKKGQTLLEENIGSAAPTGGCYLYAVLTNPFIGESLILNG
					YPAKVFSLDILGLTVESDFPLNVAVHKIKIDATASVLYSHNEALVFNTHLFO
					PIROGGLKLEKRESRELPGFSTFVFOHKKHGLMSGPCOLPCAEIMAYVITGF
					KERTVGGKLVVPVSQTPVPHIGEHOAKPLIDVDEDLGSGRAELCFRNPDISRLT
					DSIFGTIAQLARKDYSTVIAQASERFVHDQYIKPLVQAQKDFPCQASKGTDSITLV
					IDSLVALGKSTGLSTIEGQDSCEVLANTDWTIFENCETPRDLRALVWHAQSLH
					IGAQLFPAANSVLTIRVAKLPQKQNGDANMVSFYLQHLGYLSAATYKENGASAFK
					GVPVSALDSSSTYLOHLAAYASFSPLHARMCYLQFLPHIKNTNSQYNVDYVGT
					APSQMGDLQGGCPAVCINTLEFRMKDRPPVLSNKRDPYVITGTAGTYNDELTGN
					FATPFREREENGVEDAPKTYTWOLOCONTEKIASMGISRGQALTTLIVDPISFVK
					FKGIDSTVENELKPTNCKTKNNTNFRNENISVHHLQRCNRYWQAPCPVFLTYVK
					SILTVIDIDICTSCMYEODNPAVGVITSEMLMHHQYTMWNTKRGACFPKGLATGGEL
					KIVHGSNFCDLDFDTDAIGMFAPARQVYIAAMLVPTTIKIRKLITFISNTGAS
					IOAGFMKPASQDSYIVGGYKFLNAHLKHTLPSPTKSTALYLMKHLIGQTTKPIPG
					VSGEHLJELCINVKYKASSQAEELINVDVLDLTLSYAKIKINSITLRACQGOFEVAT
					LSCLSVTOCLVPAEYKPHYLGISLSPDPRYKAVGRSTYVQSTLRKQVSNGBLR
					PLITVPLVAKTKYSGNNTNVEHCANLIGFSGVDRNRLPESVPRKKNVSMMLKR
					HVTMPPLVDLVRIVIGINSGEAEVAKVSQNVLEDROPNMLPTVVALELVKHLGS
					SCASLTEDVIVYIGAYVAGDEVLSLTSLVSGAGVPADEGVAASYVTDIIDCELOF
					VGPEEPCLOGQSGVEELPSPGVPSTLVGKKRKIASLSLDL"
					6609..8696
					/note="ORF 7, processing and transport protein homolog
					(ICP18.5)"
					/codon_start=1
					/protein_id="AAB62620.1"
					/db_xref="GI:2246495"
					/translation="MAKELAAVYADVSAALMDCLVADPATLDTKSLATLTKGFS
					LHCTLLPLIRMANHCQSLLEHLLENVAALATATGCTSRKTLSPHEFSLILD
					TTCKNHRSVAFNFGMALTEKLKSLINDVEIFPKRLSSVYTCIGSSALEDGLVLR
					VGKURGISVPGPDLVVSNLPCLEQLQEVCLTNQGSLOAMLPTRASHITCPAGE
					PVGLFENELKQLGLTPESIPPTPCQSRQDEDIRSSLSLADVGDHHLFGEVTRSVL
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					









```
misc_feature      complement(6153.6296)
                    /note="match: Sts: Em: L41071"
                    19774. .20548
                    /note="Cpg Island"
misc_feature      /evidence=not_experimental
                    complement(120917.21304)
                    /note="match: GSS: Em: A2837024"
                    21687. .22022
misc_feature      /note="match: GSS: Em: A2791707"
                    24059. .24666
                    /note="match: GSS: Em: A2303512"
                    complement(126781.27451)
                    /note="match: GSS: Em: A2027075"
                    complement(126962.27413)
                    /note="match: GSS: Em: A2260334"
                    /note="match: GSS: Em: A2303774"
                    /note="match: GSS: Em: A2303774"
                    /note="match: GSS: Em: A2303774"
                    /note="match: GSS: Em: A2791347"
                    complement(33925.40432)
                    /note="match: GSS: Em: A2711476"
                    complement(33926.40442)
                    /note="match: GSS: Em: A2711491"
                    complement(44322.44838)
                    /note="match: GSS: Em: A2264680"
                    47267. .47815
                    /note="match: GSS: Em: A2753028"
                    52135
                    /note="random repeat. Forced join. Gap sized to be
                    approximately 1100bp from restriction digest data and
                    spanning pvc clone."
                    62719. .63094
                    /note="match: GSS: Em: A2259817"
                    complement(63268.63929)
                    /note="match: GSS: Em: A2087144"
                    complement(63297.63853)
                    /note="match: GSS: Em: A2032738"
                    64012. .64331
                    /note="match: GSS: Em: A2115082"
                    64026. .64425
                    /note="match: GSS: Em: A2298969"
                    complement(78311.78793)
                    /note="match: GSS: Em: A2069693"
                    complement(93998.94487)
                    /note="match: GSS: Em: A2253564"
                    94195. .94814
                    /note="match: GSS: Em: A2448780"
                    95929. .96536
                    /note="match: GSS: Em: A2394052"
                    97777. .98336
                    /note="match: GSS: Em: A2284926"
                    complement(101592.102154)
                    /note="match: GSS: Em: A2448471"
                    complement(101908.102030)
                    /note="match: GSS: Em: A2448470"
                    complement(103287.103411)
                    /note="match: GSS: Em: A2394388"
                    106254. .107006
                    /note="Cpg Island"
                    /evidence=not_experimental
                    complement(114284.114736)
                    /note="match: GSS: Em: A2041515
                    match: Sts: Em: HS28615T"
                    complement(116196.116371)
                    /note="match: GSS: Em: A2056615"
                    complement(127255.127857)
                    /note="match: GSS: Em: A2268926"
                    127968. .128455
                    /note="match: GSS: Em: A2237605"
                    complement(128635.128910)
                    /note="match: GSS: Em: A2788648"
                    143168. .143545
                    /note="match: GSS: Em: A2246254"
```

misc_feature	146147..146196	/note="Sequence from uni-directional primer reads and dGTP big dye terminator reads only."
misc_feature	complement(151206..151600)	/note="match: GSS: Em:A0930179"
misc_feature	complement(150498..152415)	/note="match: STS: Em:G071175"
misc_feature	complement(152794..153079)	/note="match: GSS: Em:A470074"
misc_feature	152905..153117	/note="match: GSS: Em:A2797680"
misc_feature	155986..160222	/note="match: GSS: Em:A284811"
misc_feature	165617..167046	/note="Cpg island"
misc_feature	/evidence=not_experimental	complement(167307..167566)
misc_feature	187090..187339	/note="match: GSS: Em:A2784357"
misc_feature	189492..189624	/note="match: GSS: Em:A2483007"
misc_feature	/note="match: GSS: Em:A2846852"	
BASE COUNT	56357 a 42439 c 41402 g 52477 t	
ORIGIN		

	Query Match	2.8%	Score 57:	DB 10:	Length 192675;	
	Best Local Similarity	44.9%	Pred.	No. 0.00029;		
	Matches 216:	Conservative	0:	Mismatches 265:	Indels	Gaps 0:
OY	305 AGGAGGAGGAGAATCTTCGTGCAGGAATTCTTAAAGAAATGACCTGGACAATGTCAGAG	364				
Db	51787 AGCAGGAGGAGCAGGAGCAGGAGCGACAGAGCACGAGGAGGAGCAGGAGCAGG	51846				
OY	365 CCCACTTTCCCAAGAAACAGGAGAAACGAGACGCCAGCTCATCTCAGCACTCTGC	424				
Db	51847 AGGACCAGGAGCAGGAGGAGGAGGAGACAGACAGCAGGAGGAGGAGGAGGAGGAGGAGG	51906				
OY	425 GGGATACGCTTGGAAMAGCAATGCTACTGTGTGATCTCTGCAGCAGGCTTGGGCAAG	484				
Db	51907 AGCAGGACGAGGAGGAGGAGGAGCAGGAGCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGG	51966				
OY	485 CCGAGATGCTGTGCTTCACACTGAAAAACGACATAGTACTTTAAGCAGCAGCAGATG	544				
Db	51967 AGCAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	52026				
OY	545 AGACCAACCAACCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	604				
Db	52027 AGCAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	52086				
OY	605 AGATTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGGTGAGGAGATGATCCGAGCATGC	664				
Db	52087 AGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	52146				
OY	665 GTGTGGGACAGTCCACCGGTGGAACAGCTGGCTGTGTACTGTGTCTCTCCAGAAAAGAGT	724				
Db	52147 AGCAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	52206				
OY	725 ACGAGATCTTAAAGAGGAGCGACGGAAGGCCCTCAGGGAGGTGGCTGACACAAGCTGAGAGAGG	784				
Db	52207 AGGACGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	52266				
OY	785 A 785					
Db	52267 A 52267					
RESULT	27					
LOCUS	AL606969	224712 bp	DNA	linear	HTG 12-DEC-2001	
DEFINITION	Mus musculus chromosome 4 clone RP23-31P21.*** SEQUENCING IN					
PROGRESS	***, in unordered pieces.					
ACCESSION	AL606969					



TITLE  
JOURNAL  
COMMENT

Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatsas,A.,  
Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehoccky,J.,  
Levine,R., Lien,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGuirk,A., McKernan,K., McPheeters,R.,  
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mienna,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testeys,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 10, 2001 this sequence version replaced gi:8139804.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information  
Center project name: L9029  
Center clone name: 228\_G\_18

## ----- Summary Statistics

Sequencing vector: Plasmid; n/a; 10% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 150458 bases at least Q40  
Consensus quality: 162515 bases at least Q30  
Consensus quality: 166983 bases at least Q20  
Insert size: 153000; agarose-fp  
Insert size: 169819; sum-of-coverage  
Quality coverage: 4.7 in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently  
consists of 43 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 13487: contig of 13487 bp in length  
\* 13488 13587: gap of 100 bp  
\* 13588 14626: contig of 1039 bp in length  
\* 14627 14726: gap of 100 bp  
\* 14727 15793: contig of 1067 bp in length  
\* 15794 15893: gap of 100 bp  
\* 15894 17057: contig of 1164 bp in length  
\* 17058 17157: gap of 100 bp  
\* 17158 18291: contig of 1134 bp in length  
\* 18292 18391: gap of 100 bp  
\* 18392 20046: contig of 1655 bp in length  
\* 20047 20146: gap of 100 bp  
\* 20147 21386: contig of 1240 bp in length  
\* 21387 21486: gap of 100 bp  
\* 21487 22760: contig of 1274 bp in length  
\* 22761 22860: gap of 100 bp  
\* 22861 23953: contig of 1093 bp in length  
\* 23954 24053: gap of 100 bp  
\* 24054 25081: contig of 1028 bp in length  
\* 25082 25181: gap of 100 bp  
\* 25182 26357: contig of 1176 bp in length  
\* 26358 26457: gap of 100 bp  
\* 26458 27574: contig of 1117 bp in length

## FEATURES

## source

27575 27674: gap of 100 bp  
27675 28756: contig of 1082 bp in length  
28757 28856: gap of 100 bp  
28857 30554: contig of 1698 bp in length  
30555 30654: gap of 100 bp  
30655 32124: contig of 1470 bp in length  
32125 32224: gap of 100 bp  
32225 33634: contig of 1410 bp in length  
33635 33734: gap of 100 bp  
33735 35520: contig of 1786 bp in length  
35521 35620: gap of 100 bp  
35621 36833: contig of 1213 bp in length  
36834 36933: gap of 100 bp  
36934 38632: contig of 1699 bp in length  
38633 38732: gap of 100 bp  
38733 40693: contig of 1961 bp in length  
40694 40793: gap of 100 bp  
40794 41899: contig of 1106 bp in length  
41900 41999: gap of 100 bp  
42000 43910: contig of 1911 bp in length  
43911 44010: gap of 100 bp  
44011 45470: contig of 1460 bp in length  
45471 45570: gap of 100 bp  
45571 46993: contig of 1423 bp in length  
46994 47093: gap of 100 bp  
47094 48770: contig of 1677 bp in length  
48771 48870: gap of 100 bp  
48871 51099: contig of 2223 bp in length  
51100 51199: gap of 100 bp  
51200 53719: contig of 2520 bp in length  
53720 53819: gap of 100 bp  
53820 56012: contig of 2193 bp in length  
56013 56112: gap of 100 bp  
56113 57226: contig of 1114 bp in length  
57227 57326: gap of 100 bp  
57327 77403: contig of 20077 bp in length  
77404 77503: gap of 100 bp  
77504 79994: contig of 2491 bp in length  
79995 80094: gap of 100 bp  
80095 82923: contig of 2823 bp in length  
82924 83023: gap of 100 bp  
83024 86207: contig of 3184 bp in length  
86208 86307: gap of 100 bp  
86308 90457: contig of 4150 bp in length  
90458 90557: gap of 100 bp  
90558 96908: contig of 6351 bp in length  
96909 97008: gap of 100 bp  
97009 102861: contig of 5853 bp in length  
102862 102961: gap of 100 bp  
102962 107881: contig of 4920 bp in length  
107882 107981: gap of 100 bp  
107982 113343: contig of 5562 bp in length  
113344 113643: gap of 100 bp  
113644 118163: contig of 4520 bp in length  
118164 118263: gap of 100 bp  
118264 135271: contig of 17008 bp in length  
135272 135371: gap of 100 bp  
135372 151730: contig of 16359 bp in length  
151731 151830: gap of 100 bp  
151831 171146: contig of 19316 bp in length  
171147 171246: gap of 100 bp  
171247 174019: contig of 2773 bp in length.

Location/Qualifiers  
1. 174019

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP11-228G18"

/clone\_11b="RPC1-11 Human Male BAC"

1. 13487

/note="assembly\_fragment"

clone\_end:SP6

## misc\_feature

```
misc_feature      13588.14626      vector_side:left"
misc_feature      14727.15793      /note="assembly_fragment"
misc_feature      15894.17057      /note="assembly_fragment"
misc_feature      17158.18291      /note="assembly_fragment"
misc_feature      18392.20046      /note="assembly_fragment"
misc_feature      20147.21386      /note="assembly_fragment"
misc_feature      21487.22760      /note="assembly_fragment"
misc_feature      22861.23953      /note="assembly_fragment"
misc_feature      24054.25081      /note="assembly_fragment"
misc_feature      25182.26357      /note="assembly_fragment"
misc_feature      26458.27574      /note="assembly_fragment"
misc_feature      27675.28756      /note="assembly_fragment"
misc_feature      28857.30554      /note="assembly_fragment"
misc_feature      30655.32124      /note="assembly_fragment"
misc_feature      32225.33634      /note="assembly_fragment"
misc_feature      33735.35520      /note="assembly_fragment"
```

Query Match	2.88;	Score 56.2;	DB 2;	Length 174019;
Best Local Similarity	45.08;	Pred. No. 0.00048;		
Matches 250;	Conservative 0;	Mismatches 303;	Indels 2;	Gaps 1

OY	376	CAAAAAGCAGAGCAAAAACGACACAGCCAGGTCATCATCATCTCGGGGATCGGCTG	435
Db	66805	CAGAAAGCAGAGAGCAAAATATGTGGGACGACAGGAAGAAAGAGCTGGCAGCAGAGAGCTT	66864
OY	436	GAGAAAGCCAAATGCTACTGTGGTATCTTCGACGACGCCCTTGGGCAAGGCCAGATGCTG	495
Db	66855	CCGGAACCAAGAGCAAGCTGTGGGACACGAGAGAAAGATTCAGAGACGAGGAGGAAAGATA	669224
OY	436	TGCTCCACACTGTA AAAAGCAGATGAATGACTTAGAGCAGCAGCAGAGATGAGACCAACAA	555
Db	66925	TGGGAGCAGAGAGAAAGATATACGGGACACGAGAGAGATGTGGGGCAGGAGAAAGATG	669804
OY	556	GCACAAGAGAGAGCGGGCGGCTCAGAGACCAAGATGAAAGCATTGAGACATTTAGCTT	615
Db	66985	TGGCGGCGAGGAAMAATGCGG--GAGCAGGAAGATGTGGAGACAGAGAGAGCGCTGCAG	67042
OY	616	CTACTCCAGAGCAGCTCCCTGAGGAGGAGAGATGATCCGAGACATGGGCTGTGGGACAG	675
Db	67043	GAGCAGGAGAAACAAATGTGGGAGCAGAGGAGGAGAAAGATCGGGATCAGGAGCAAGATG	67102
OY	676	TCACGGGTGGAACACTGGCTGTACTGTGTCTCTCAAAAAGATGACAGATCTA	735
Db	67103	TGGGACACAGAGAGAGATGTGGGAGCAGACGAGAGGCTCGGGAGAAAGAGAGAGACA	67162
OY	736	AAAGAGGACCGAAGGCCCTCAGGGGAGGTGCTGACAAAGCTAGGAGAGATTGTTTTCC	795
Db	67163	ATGCGGAGGACAGAAAGATGTGGCAGCAGTGTGGAAGATGTCGGGAGAGGAAGAACG	672222
OY	796	TTCACAGACAACTTCAGACAGTCTACTCTGATTTGGATCAGGCCAAGTTAGAACTGAAG	855
Db	67223	CAGGACGAGGAGAAAGACACATGGGACCAAGAGGAGAAAGATCCGAGAGGAGAGAGATG	672822
OY	856	TCAGCCCAAGAGCACTTACAGAGTGTCTACAAAGAAATCATATGACCTGAAAAAGAGCTA	915
Db	67283	CGGGAGCGGGAGAAAGATGCGGGAGGAGGAGGAGAGATCATCTCGGAGCAGGAGAGGAGAG	67342

oy	916	ACGATCGTCAGGAA	930
db	67343	ATGCAGGAGCAGGAA	67357
RESULT	29		
AC073446/c			
LOCUS			
DEFINITION			
AC073446			
AC073446			
AC073446.7			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			

185497 bp DNA linear HTG 11-OCT-2001  
 Homo sapiens chromosome 15 clone RP11-757E13 map 15, \*\*\* SEQUENCING  
 IN PROGRESS \*\*\*, 5 unordered pieces.  
 AC073446  
 AC073446.7 GI:16041356  
 HTG: HTGS\_PHASE1; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 185497)  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 Homo sapiens chromosome 15, clone RP11-757E13  
 Unpublished  
 2 (bases 1 to 185497)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
 Boguslavsky,I., Bouckgalter,B., Brown,A., Burkett,G.,  
 Campoliano,A., Castle,A., Chepel,Y., Colangelo,M., Collins,S.,  
 Collamore,A., Cooke,P., Dearliano,K., Dewar,K., Diaz,J.S.,  
 Dodge,S., Domini,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,  
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
 Girard-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
 Klein,J., Lacombe,K., Lamazeres,R., Landers,T., Lehoczy,J.,  
 Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., Mcpheeters,R.,  
 Meldrum,T., Menes,L., Minova,T., Miranda,C., Mleaga,V., Morrow,J.,  
 Murphy,T., Naylor,T., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 O'Neill,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N.,  
 Pisan,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
 Stange-Thoman,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Testa,S., Theodore,J., Tirrell,A., Travers,A., Triggillo,D.,  
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,D., Zimmer,A. and Zody,W.  
 Direct Submission  
 Submitted (18-JUN-2000) Whitehead Institute/MIT Center for Genome  
 Research 320 Charles Street, Cambridge, MA 02141, USA  
 On Oct 11, 2001 this sequence version replaced gi:15718563.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center -----  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www.seg.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information -----  
 Center project name: L10251  
 Center clone name: 757\_E\_13  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 5 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 33974: contig of 33974 bp in length  
 \* 33975 34074: gap of 100 bp  
 \* 34075 115332: contig of 81258 bp in length  
 \* 115333 115432: gap of 100 bp  
 \* 115433 117817: contig of 2385 bp in length



```

Db      34963  GGAGAAAGGGAAGAAGACAGACAGACAAGCAAGAAGAAAGCAAGAACAGACAGAACAGAA 35022
Oy      870   CTTACAGAGTGTCTGACAAAGAAATCATGAGCCTGTAAAAAAGACTPACGATGCTTCAGGA 929
          |||         | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      35023  GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 35082
Oy      930   AACCTTGAA 938
          | | | | |
Db      35083  GAAAGAAAGAA 35091

RESULT  31
AL645746
LOCUS    AL645746                282611 bp     DNA             linear       HTG-12-JAN-2002Z
DEFINITION Mus musculus chromosome 13 clone RP23-153B6, *** SEQUENCING IN
PROGRESS ***, in unordered pieces.
ACCESSION AL645746
VERSION   AL645746.2 GI:17148444
KEYWORDS  HTG, HTGS_PHASE1.
SOURCE    house mouse.
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS   Sims,S.
TITLE      Direct Submission
JOURNAL    Submitted (11-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
           Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
           hinxton@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
           On Nov 28, 2001 this sequence version replaced gi:17017892.

COMMENT
           genome Center
           Center: UK Medical Research Council
           Center code: UK-MRC
           Web site: http://mrcseq.har.mrc.ac.uk
           Contact: mouse@har.mrc.ac.uk
           ----- Project Information
           Center project name: Bmi13B6
           ----- Summary Statistics
           Assembly Program: XGAP4; Version 4.5
           Sequencing vector: Plasmid; 108752; 100% of reads
           Chemistry: Dye-terminator Big Dye; 100% of reads
           Consensus quality: 279809 bases at least Q40
           Consensus quality: 280516 bases at least Q30
           Consensus quality: 281062 bases at least Q20
           Insert size: 281711; sum-of-contigs
           Insert size: 190968; 4.7% error; agarose-fp
           Quality coverage: 12.39x in Q20 bases; sum-of-contigs Quality
           coverage: 18.52x in Q20 bases; agarose-fp
           -----
           * NOTE: This is a 'working draft' sequence.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.

FEATURES
Source
location/Qualifiers
1..282611
   /organism="Mus musculus"
   /db_xref="taxon:10090"
   /chromosome="13"
   /clone="RP23-153B6"
   /clone_1fb="RPCR-23"
1..30613
   /note="assembly_fragment:06533
fragment_chain:1"
30714..46652
   /note="assembly_fragment:02685
fragment_chain:1"
46753..51050
   /note="assembly_fragment:08628
fragment_chain:1"
51151..57360
   /note="assembly_fragment:06992
fragment_chain:1"
57461..61884

```

		/note="assembly_fragment:03821 fragment_chain:1"
misc_feature		/note="assembly_fragment:05324 fragment_chain:1" 71985..71188
misc_feature		/note="assembly_fragment:05094 fragment_chain:2" 71289..219059
misc_feature		/note="assembly_fragment:06258 fragment_chain:2" 250588..257101
misc_feature		/note="assembly_fragment:05089" 257202..282611
misc_feature		/note="assembly_fragment:06688" BASE COUNT    80227 a    60936 c    60357 g    80189 t         902 others
ORIGIN		
Query Match	2.8%	Score 56; DB 2; Length 282611;
Best Local Similarity	45.6%;	Pred. No. 0.00059;
Matches 197; Conservative	0;	Mismatches 235; Indels 0; Gaps 0;
QY	507	GAAAAAGCATGAACTACTTAGACGACGACGATGGACCACAACACCAAGAGA 566
Db	219182	GAAAGGAAGGAGGAGAGGAGCAGAACAGGAGCGAGGAGGAAAGAGAGAGAGAGA 219241
QY	567	GGCCGCCGCCTCACAGCAGCATGTGAACCATGGACGATTGAGCTTCTACTCCAGAG 626
Db	219242	GAAAGGAGAAAGA 219301
QY	627	CCAGCTCCCTGAGGTGAGTGAGAGATGATGCCAGATCGGTGTGGACACTCAGCGGTGA 686
Db	219302	GGAGGAGAAAGA 219361
QY	687	AACAGTCGCTGTGTCTACTGTGTCTCTCAAAGAAAGACTACGGAATCTAAAAGAGCCAG 746
Db	219362	GGAGGAGAAAGGAGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA 219421
QY	747	GAAAGCGCTCAGGAGGAGGTGCTGACAGCTGAGAAAGATTTGTTCCTCCAGAAAGCA 806
Db	219422	GAAAGGAGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA 219481
QY	807	GTTGACAGACAGTCTACTCTGAAATTGGATCAGCCCAAGTTAAGTAAGTCAAGTCAAGCCAGAA 866
Db	219482	GGAGGAGAAAGA 219541
QY	867	GCACCTTACAGACTGCTGACACAGAAATCATGAGCCTGAAAAAGAAAGCTTACCATCTGCA 926
Db	219542	GGAGAGAAAGGAGAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGA 219601
QY	927	GGAACCTTGAA 938
Ddb	219602	GAAAGAGAGAGAA 219613
RESULT 32		
AFI48805/c		
LOCUS	AFI48805	28559 bp DNA linear VRL 02-AUG-1999
DEFINITION	Kaposi's sarcoma-associated herpesvirus ORF 68 gene, partial cds; and ORF 69, Kaposin, v-FLIP, v-cyclin, latent nuclear antigen, ORF k14, v-GPCR, putative phospholipid synthase, complete cds.	
ACCESSION	AFI48805	
VERSION	AFI48805.1	GI:5669888
KEYWORDS	Kaposi's sarcoma-associated herpesvirus - Human herpesvirus 8. Human herpesvirus 8 Viruses: dsDNA viruses, no RNA stage: Herpesviridae; Gammaherpesvirinae: Rhadinovirus.	
SOURCE	HUMAN	
ORGANISM	(bases 1 to 28559)	
REFERENCE	Glenn,M., Rainbow,L., Aurd,F., Davison,A. and Schulz,T.F. Identification of a spliced gene from Kaposi's sarcoma-associated	
TITLE		



FEATURES	source
JOURNAL	herpesvirus encoding a protein with similarities to latent membrane proteins 1 and 2A of Epstein-Barr virus
MEDLINE	J. Virol. 73 (8), 6953-6963 (1999)
REFERENCE	2 (bases 1 to 28559)
AUTHORS	Glenn, M.A., Rainbow, L., Aurade, F., Davison, A.J. and Scholz, T.F.
TITLE	Submitted (15-APR-1999) MRC Virology Unit, Church Street, Glasgow G11 5UR, UK
JOURNAL	Location/Qualifiers
source	1. .28559 /organism="Human herpesvirus 8" /isolate="GK18" /db_xref="taxon:37296" /note="right end of the unique region plus 8 complete copies and 1 partial copy of the terminal repeat (TR)" <1. .2105 /gene="ORF 68" <1. .1102 /gene="ORF 68" /function="involved in DNA packaging" /note="similar to HSV-1 UL32, HCMV UL52 and EBV BFLF1" /codon_start=2 /evidence=not_experimental /protein_id="AAd6496.1" /db_xref="GI:5669889" /translation="IMHLHPEEEATYGLFGPGRMGIDLIOLHFFVOCKEKTAAEKI LGISLLOKSEFIRGLMTGTITGNPCFKTSPRTDKEEANGPTGCCOITDTTAPAS GIPRLARATCGASRPTKPTSLIRPLALDIWTSSSLDEPRRLIASMSSELKQVYASH DFFPSPLADTSQGPCILMHTTGLKRNKGASVCLLCECIAHPEAPKALQTLQCEVH MGI,ENNVKLLVDRIAEFLVLDNFPAPVYSDPLRLRLINGCTPOEIHKLHLPDCLALNA KVSSEVDLFLRPREOEYKLRASAAOGLDANTLPCEVQVQLVFLFKGLQNARVKG TTSDIIRLELTAQLKRRRLDLAHPQSHTLYA" 1135. .2105 /gene="ORF 69" 1135. .2043 /gene="ORF 69" /note="similar to HSV-1 UL31, HCMV UL53 and EBV BFLF2" /codon_start=1 /evidence=not_experimental /protein_id="AAd6497.1" /db_xref="GI:5669890" /translation="MPKSVSSHISLATSTGRSGPRDIRCLSLRSVPPGARSASVYS SKHNGLRKFTISDKVVFSLSHRELGVDFLRKETP ICTSKTYMLPLDSTVPGRC VSLPFGGSSNMGFOCALCPETNPVAOGRPOTMGDALKKNELCSVALAEYHHA DKVIOHKTEVYLSLSHGMADVROSEFLQGLLYANLVKTRGHDPLEPTNGMGLTMC ILFTRALHGETRALRLMDNLPYRKISADCCRSQSYVVKFVPTHPASTIAVQVHTTC EAAVALDCTDMDRDIKGTALNAL" 2100. .2105 /gene="ORF 69" /note="possible signal for ORFs 68 and 69" /evidence=not_experimental complement(2126) /gene="ORF K12" /evidence=experimental complement(2126. .2824) /gene="ORF K12" /product="10.7" /evidence=experimental complement(2126. .2824) /gene="ORF K12" complement(2139. .2144) /gene="ORF K12" /evidence=experimental complement(2616. .2798) /gene="ORF K12" /note="transforming gene" /codon_start=1 /evidence=experimental /product="Kaposin" /protein_id="AAd6499.1" /db_xref="GI:5669892" /translation="MDGRGIFVAVVHVDVLLNGMRMRGAIPPLYVCLLAISVPPSGS
gene	
CDS	
gene	
CDS	
polyA_signal	
polyA_site	
mRNA	
gene	
polyA_signal	
CDS	

[illegible]

```

repeat_region
gene
CDS
/note="complex array of repeats"
/rpt_type=direct
12624..15262
/gene="ORF K14"
12624..13670
/gene="ORF K14"
/note="similar to OX-2 membrane glycoprotein"
/codon_start=1
/evidence=not_experimental
/protein_id="AADA6502.1"
/db_xref="GI:5669895"
/translation="MHHFPPCGRRVVEGVSSYFLIGAPGPTAKTKEGVSALON
LPPVLPVAGGGSVRRPVVCAAPPTTSSASGSSSLFISLPKAFIWLALLGAVGA
RVGPMKSGAALVLAITPRADIVSYVKQRLPQPVAVATSHSYGVVDTQYRKNM
ITCGMLNSTMVLIHLAVDDEGCYLIENSGVSCYCTALEVTSPTSHOVNSTE
DADVITCLATGRRPPPTVMAAPMNNASSTQEQFDISGLITVAMTVLIPGDDTPEE
GICITLGNESISIPASIOGPLAHDLPAAGCTAGVATITVGLFGIPALHNCRKQGG
ASPPSDMDPLSTQ"
14111..15262
/gene="ORF 74"
14111..15139
/gene="ORF 74"
/note="G protein coupled receptor; similar to IL-8
receptor; transforming protein"
/codon_start=1
/evidence=experimental
/product="V-GPCR"
/protein_id="AADA6503.1"
/db_xref="GI:5669896"
/translation="MAAEDFLTFLDDDESMETLNMGSYDYSNFSLEYSVCEMTLV
VPYRMNGLISLILNVLNGLVTYIFCKHRSRAGAILDILIGICVNSCLISLSLA
EVLNLFEPNIIISTGLCRLEIFETLYLYLIDIFSVCVSLVRKLLVAITSWPKQSL
Query Match 2.8%; Score 55.8; DB 14; Length 28559;
Best Local Similarity 43.4%; Pred. No. 0.00049;
Matches 255; Conservative 0; Mismatches 332; Indels 0; Gaps 0;
OY 304 CAGCAGGAGGAGATGCTTGATCGACAATTCCTTAAAGATGAACCTGGACAATGTCAGA 363
Db 10156 CAGCAGGATGAGCAGACAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGCAGAGAT 10097
OY 364 GCCCGCTTTOCCCAAAAGACAAGAGAAAGACAAGCAGCAGCAGCAGATCATATCGACACTGTG 423
Db 10096 GAACGAGGACGACAGGAGAGCAGGAGGAGCAGGAGGAGGAGCAGCAGAGATGACACAGCAG 10037
OY 424 CGGAGTACGCTGGGAAGACGCAATCTCTGTGTATCTCTGCAGCAGGAGCCTTGGGCAAG 483
Db 10036 CAGGATGAGCAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGCAGCAGATGAGCAG 9977
OY 484 GCCGAGATCTGTGCTCCACACTGAAAAAGCAGATGAAGTACTTAGACAGCAGCAGAGAT 543
Db 9976 CAGCAGGATGAGCAGCAGCAGCAGCAGATGAACAGGAGCAGCAGAGAGAGCAGCAGCAGCAG 9917
OY 544 GAGACCAACAACAACAAGAGAGAGGCGCGCGCTCAGAGACCAAGATGAACAACATGAGAG 603
Db 9916 GAGGAGCAGGAGCAGCAGAGAGAGCAGGAGCAGCAGGAGGAGCAGGAGCAGGAGTTAGAG 9857
OY 604 CAGATTGAGCTTCTACTCCAGAGCCAGCTCCCTGAGTGGAGGAGAGATGATCCGAGACATG 663
Db 9856 GAGCAGGAGCAGGAGATTAGAGAGCAGCAGCAGCAGCAGGATTTAGAGGAGCAGCAGCAGAGCTTA 9797
OY 664 GGTGTGGCAGATCAGCGGTGGAACAGCTGGCTGTGTACTTGTCTCTCAAGAAAGAG 723
Db 9796 GAGGAGCAGGAGCAGGAGATTAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 9737
OY 724 TACGAGAACTTAAAGAGCAGCAGGAGCCCTCAGGGAGGTGGCTGACACAGCTGAGAGAG 783

```

[illegible]

```
Query Match      2.8%; Score 55.4; DB 14; Length 3127;
Best Local Similarity 45.6%; Pred. No. 0.00047;
Matches 194; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

QY 505 CTGAAAAGCAGATGAACTTACTTAGACAGACAGCAGATGAGACCAAAACAGCACAAG 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1738 CACACAGGATGAGCAGCAGCAGATGAGCAGCAGCAGCAGATGACAGGAGCAGCAGAG 1797
QY 565 GAGCGCGGCCGCTCAGAGCAGCAAGATGAGACCATGAGCAGATTGAGCTTTACTTCCAG 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1798 GACAGCAGCAGCAGCAGAGCAGCAGCAGCAGCAGTGTAGAGCAGCAGCAGCAGATTAGAG 1857
QY 625 AGCCAGCTCCCTGAGGAGGAGATGATCGAGACATGGGTGGAGCAGTCAAGCGGCTG 684
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1858 GACAGCAGCAGCAGATTTAGAGGAGCAGCAGCAGAGTGTAGAGCAGCAGCAGAGATT 1917
QY 685 GAACAGCTGCGCTGTACTGTCTCTCTCAAGAAAGATGACGAGAATCTTAAAGAGCGA 744
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1918 GAGGAGCAGGAGCAGCAGGATTAGAGGAGCAGCAGCAGAGTGTAGAGCAGCAGCAGAG 1977
QY 745 CGGAGGCTTCAGGCGCAGTGCCTGACACAGCTGAGAGGATTTGTTTCTCCAGAAAGC 804
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1978 TTAGAGAGCAGCAGCAGCAGGATTAGATGACAGCAGCAGAGGAGTTAGAGCAGAGCAG 2037
QY 805 AAGTTGACAGACGCTTACTCTGATTCGATCAGGCCAAGTTAGAACTGAAGTCAGCCAG 864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2038 GAGCTTAGAGAGCAGCAGGAGTTAGAGCAGCAGCAGAGAGGAGTTAGAGGAGCAGAG 2097
QY 865 AAGGACTTTACAGATGCTGACAAAGAAATCATGAGCCTGAAAAAGACTAACGATCTG 924
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2098 CAGAGGTTAAGAGCAGCAGCAGCAGGAGCTTAGAGAGCAGCAGCAGAGATTAGAGAGCAG 2157
QY 925 CAGCA 929
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2158 GAGCA 2162

RESULT 34
AC098712 207683 bp DNA linear HTG 14-NOV-2001
LOCUS Mus musculus chromosome UNK clone RP23-2A21, WORKING DRAFT
DEFINITION
AC098712 AC098712
SEQUENCE 5 unordered pieces.
AC098712.2 GI:16924164
AC098712.2 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 207683)
McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 207683)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (31-OCT-2001) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Nov 14, 2001 this sequence version replaced gi:16554409.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_BA0002A21
----- Summary Statistics -----
Sequencing vector: MJ3; %
Sequencing vector: plasmid; %
```

```
Chemistry: Dye-primer ET; % of reads
Assembly: Dye-terminator Big Dye; % of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 205930 bases at least Q40
Consensus quality: 206594 bases at least Q30
Consensus quality: 207083 bases at least Q20
Insert size: -; agarose-fp
Quality coverage: 25.41 in Q20 bases; agarose-fp
Quality coverage: 24.42 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1240: contig of 1240 bp in length
* 1241 1340: gap of unknown length
* 1341 14229: contig of 12889 bp in length
* 14230 14329: gap of unknown length
* 14330 39801: contig of 25472 bp in length
* 39802 71680: gap of unknown length
* 71681 71781: contig of 31779 bp in length
* 71781 207683: gap of unknown length
* 207683: contig of 135903 bp in length.
FEATURES
source
1..207683
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-2A21"
1..1240
/feature="assembly_name:Contig14"
1341..14229
/feature="assembly_name:Contig17"
14330..39801
/feature="assembly_name:Contig38"
39902..71680
/feature="assembly_name:Contig39"
71781..207683
/feature="assembly_name:Contig40"
BASE COUNT 62336 a 44191 c 44243 g 56460 t 453 others
ORIGIN

Query Match      2.8%; Score 55.4; DB 2; Length 207683;
Best Local Similarity 44.7%; Pred. No. 0.00083;
Matches 215; Conservative 0; Mismatches 266; Indels 0; Gaps 0;

QY 305 AGGAGAGGAGAAATGCTTGATCGAATTTTAAAGATGAACTGACATGTGACAG 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 186130 AGGAGAGGAGAGAGCAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186189
QY 365 CCCAGCTTTCCCAAGAAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 186190 AGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186249
QY 425 GGGATTACGCTGGAGAAAGCAATGCTTGTGATCTGTGATCTGTGATCTGTGATCTGTG 484
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 186250 AGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186309
QY 485 CCGAGATGCTGTGCTCCACATGAAAAGCAGATGAGTAACTTAGAGCAGCAGCAGAGATG 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 186310 AGGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186369
QY 545 AGACCAACAAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 186370 AGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186429
QY 605 AGATTGAGCTTACTCTCAAGCAGCAGCTCCCTGAGGTGAGAGAGATGATCGAGACATGG 664
```

	DB	186430	ACGAGGAGAGAGAAGGACGACGAGCAGAGAGACGAGAGAGCAGCAGGACGACGAGAACG	1864189
OY	665	GTCGTGGACAGCTCACCGGTGGCAACACTGCTGTGTACTGTGTCTTCAAGAAAGCT	724	
DB	186490	ACCAGAGACGACGAGAGAGCAGGAGGAGCAGAGAGACGCTGGAGAGAGACGCCGAGAGGAGAC	186549	
OY	725	ACGGAATTAATAAGAGGCACGAGAGCCCTCAGGGAGGTGGCTGTGACAAGCTGACGAGAAGC	784	
DB	186550	AGGAGAGACGAGAGAAGAGGAGGAGGAGGAGCAGCAGAGAGCAGCAGGACGACGAGAGC	186609	
OY	785 A	785		
DB	186610	A	1866610	
RESULT	35			
ACOL1553				
LOCUS	ACOL1553	35848 bp	DNA	linear
DEFINITION	Homo sapiens chromosome 19 clone LTNLR-277D11, WORKING DRAFT			
SEQUENCE	3 ordered pieces.			
ACOL1553				
ACOL1553	3	GI:7711540		
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEPTN.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	DOE Joint Genome Institute.			
JOURNAL	Sequencing of Human Chromosome 19			
REFERENCE	Unpublished			
AUTHORS	2 (bases 1 to 35848)			
TITLE	DOE Joint Genome Institute.			
JOURNAL	Direct Submission			
AUTHORS	Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint			
JOURNAL	Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA			
COMMENT	On May 6, 2000 this sequence version replaced gi:7690187. -----Genome Center Center: Joint Genome Institute Center Code: JGI Web site: http://www.jgi.doe.gov ----- Project Information Center Project Name: 26813, R31543 Center clone name: LLNL-R_277D11 ----- Summary Statistics Consensus quality: 35124 bases at least Q40 Consensus quality: 35398 bases at least Q30 Consensus quality: 35534 bases at least Q20 Estimated insert size: 32650; agarose-fp estimation Estimated insert size: 35748; sun-of-contigs estimation Quality coverage: 10.55 in Q20 bases; agarose-fp estimation Quality coverage: 9.64 in Q20 bases; sum-of-contigs estimation. NOTE: This is a 'working draft' sequence. It currently * consists of 3 contigs. Gaps between the contigs * are represented as runs of N. The order of the pieces * is believed to be correct as given, however the sizes * of the gaps between them are based on estimates that have * provided by the submittor. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved. 1 * 18613: contig of 18613 bp in length * 18614 * 18713: gap of unknown length * 18714 * 19277: contig of 564 bp in length * 19278 * 19377: gap of unknown length * 19378 * 35848: contig of 16471 bp in length. Location/Qualifiers 1..35848 /organism="Homo sapiens" /db_xref="taxon:9606"			
FEATURES				
SOURCE				

```

/chromosome="19"
/clone="L1NLNR-277D11"
/clone.lib="Lawrence Livermore human cosmid library L1NLNR
BASE COUNT      7758 a 11001 c 10770 g 6119 t      200 others
ORIGIN

Query Match      2.7%; Score 55; DB 2; Length 35848;
Best Local Similarity 49.8%; Pred. No. 0.0085;
Matches 139; Conservative 0; Mismatches 140; Indels 0; Gaps 0.

QY  507  GAAAAAGCAGATGAGTAGTACTTAGACAGACGAGGATGACCAACAACGACAAGAGA 566
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  19523 GGAGGAAACGGAGAGAGAGAGAGAGAGAAACAGGAGGAGGAGAGAGAGAGAGAGA 19582
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  567  GCGCGCGCGCGCTCAGAGAGCAGCAAGATGAACGACATGAGAGCATTTGACTTCTACTCCAGAG 626
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  19583 GGAGGAGGAGGAGAAACAGCAGCAGAGAGAGAGAGGCGACACAGGAGGAGGAGAGGAGGAAGAACACAGAGAGA 19642
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  627  CCAGCTCCCTGAGTGAGTGAGAGAGATGATCCGAGACATGGGTGTGGGACACTGACGGGTGA 686
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  19643 GGAGGAGGAGAAACAGGAGAGAGAGAGAGGAGAAACAGGAGGAGGAGGAGAGGAGGAACAGAGAGAGAGAGA 19702
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  687  ACAGCTGGCTGTGTACTGTGTGTCTCTCAAGAAAGATTACGGAATCTTAAAGAGGCAACG 746
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  19703 GGAAACAGGAGGAGGAGGAGAGAGAGAGGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGAGAGAGA 19762
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  747  GAAAGCCTCAGGGGAGGAGGTGGCTGACAAAGCTGAGGAAGGA 785
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  19763 GGAGGAGGAGAAACAGGAGAGAGAGAGAGGAGAAACAGGAGAGAGA 19801

```

Query Match	Best Local Similarity	Score	DB 2:	Length	110000:
Matches 225: Conservative	0:	Mismatches 285:	Indels	0:	Gaps 0:
Continuation (3 of 4) of AC091454 from base 200001 (AC091454 Mus musculus chromosome					
Sequence split into 4 fragments					
Fragment Name	Begin	End	LOCUS	AC091454	Accession AC091454
AC091454_0	1	110000			
AC091454_1	100001	210000			
AC091454_2	200001	310000			
AC091454_3	300001	394326			
Query Match 2.7%: Score 54: DB 2: Length 110000:					
Best Local Similarity 44.1%: Pred. No. 0.0019:					
Matches 225: Conservative 0: Mismatches 285: Indels 0: Gaps 0:					
QY 305	AGGAGGAGAGAATGTCCTGGATCGAATTCCTTAAAGATGAACATGACATGTCAGAG	364			
Db 69201	AGGACAGAGAGACGAGGAGGAGGAGCAGCAGCAGAGAGAGGAGGAGGAGGAGGAGGAGG	69142			
QY 365	CCCACCTTTCCAGAAAGACAGGAGGAGAAACGAGACAGCCAGTCATCTATCTGAC	424			
Db 69141	AGGACACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	69082			
QY 425	GGGATACGCTGGAAACACCAATGCTACTGTGATCTCTGTCAGCAGGCTTTGGGACAGG	484			
Db 69081	AGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	69022			
QY 485	CCGAGATGCTGTGCTCCACACTGAAAAAGCAGATGAATGACTTTAGACACGACAGAGATG	544			
Db 69021	AGG	68962			
QY 545	AGACCAACCAACGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	604			
Db 68961	AGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	68902			
QY 605	AGATTGACCTTCTACTCCAGACGACCTCTCTGAGGTGAGGAGGAGATGATCCGAGACATGG	664			
Db 68901	AGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	68842			
QY 665	GTGTGGGACAGTCACCGGTGGAAACAGCTGGCTGTGACTGTGTCTTCACAAAGAAAGCT	724			

```
Db 68841 AGCAGAGCGACGAGCAGAGAGATCAGGAGAGCAGCAGCAGAGAGAGG 68782
Qy 725 AGCAGATCAAAAAGGCGCAGGAGGCCCTCAGGAGGTGCTGCACACCTGAGAGG 784
Db 68781 AGAGGAGCGAGGAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGCAGCACC 68722
Qy 785 ATTTGTTTCTCCAGACGACAGTTCGACA 814
Db 68721 ATGAGGAGGGGACAGAGACAGACAGAGA 68692

RESULT 37
AC007461      180385 bp      DNA      linear      PRI 04-JUN-1999
LOCUS        Homo sapiens chromosome 17, clone 84_E_24, complete sequence.
AC007461
AC007461.8   GI:4996920
DEFINITION   HTG.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM     Homo sapiens
              Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE    1 (bases 1 to 180385)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE        Homo sapiens chromosome 17, clone 84_E_24
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 180385)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
              Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
              Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
              Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,
              Donejan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,
              Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grunt,G.,
              Hagos,B., Heathord,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
              Karatas,A., Lehoccky,J., Lieu,C., Locke,K., Macdonald,P.,
              Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
              Meldrum,J., Molla,M., Morris,M., Morrow,J., Mychalecky,J.,
              Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
              Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
              Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
              Testaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
              Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
TITLE        Direct Submission
JOURNAL      Submitted (04-MAY-1999) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE    3 (bases 1 to 180385)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
              Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
              Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
              Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,
              Donejan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,
              Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grunt,G.,
              Hagos,B., Heathord,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
              Karatas,A., Lehoccky,J., Lieu,C., Locke,K., Macdonald,P.,
              Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
              Meldrum,J., Molla,M., Morris,M., Morrow,J., Mychalecky,J.,
              Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
              Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
              Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
              Testaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
              Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
TITLE        Direct Submission
JOURNAL      Submitted (04-JUN-1999) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT      On Jun 4, 1999 this sequence version replaced gi:4966400.
              All repeats were identified using RepeatMasker: Smit, A.F.A. &
              Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html

Only 180385 base pairs from the middle of this clone are being
submitted. The remainder overlaps either accession number AC005152
(WICGR project l335) or accession number AC005243 (WICGR project
```

```
FEATURES             L343).
source
Location/Qualifiers
1..180385
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="84_E_24"
/clone_lib="Alan Buckler -- per comm"
446..753
/rpt_family="AluSc"
2375..2419
/rpt_family="(TTTGG)n"
complement(2985..3103)
/rpt_family="L2"
complement(3146..3352)
/rpt_family="L2"
complement(3531..3639)
/rpt_family="L2"
complement(3798..4394)
/rpt_family="L2"
5531..5631
/rpt_family="MER81"
5648..5674
/rpt_family="(CAAAA)n"
5812..6020
/rpt_family="LIMC4"
6180..6224
/rpt_family="AT_rich"
6303..6611
/rpt_family="AluSp"
7378..7618
/rpt_family="MTR"
9593..9623
/rpt_family="(GGGA)n"
complement(10666..10923)
/rpt_family="AluSq"
10958..11088
/rpt_family="MTR"
11752..11914
/rpt_family="L2"
complement(11915..12219)
/rpt_family="AluSx"
12220..12347
/rpt_family="L2"
13110..13182
/rpt_family="GA-rich"
13526..13564
/rpt_family="(TCC)n"
13575..13660
/rpt_family="(TR)n"
complement(14380..14658)
/rpt_family="AluJo"
14742..14830
/rpt_family="(TA)n"
14860..14967
/rpt_family="(TA)n"
complement(14977..15122)
/rpt_family="L2"
complement(15292..15385)
/rpt_family="L2"
15672..16781
/rpt_family="L2"
16802..16846
/rpt_family="(TGAA)n"
17225..17245
/rpt_family="(A)n"
18411..18720
/rpt_family="AluJb"
complement(20715..20788)
/rpt_family="MTR1J2"
complement(20822..20916)
/rpt_family="MTR1J2"
```

```

repeat_region 20956. .21137
/rpt_family="MIR"
repeat_region 21830. .21872
/rpt_family="AT_rich"
repeat_region 21957. .22257
/rpt_family="Alusx"
repeat_region complement(22438. .22538)
/rpt_family="MLT1J"
repeat_region complement(23375. .23667)
/rpt_family="Aluy"
repeat_region 24553. .24858
/rpt_family="Aluy"
repeat_region complement(25021. .25234)
/rpt_family="MIR"
repeat_region 26957. .27261
/rpt_family="Aluy"
repeat_region complement(27524. .27669)
/rpt_family="MLT1B"
repeat_region complement(27670. .27964)
/rpt_family="AlusP"
repeat_region complement(27965. .28210)
/rpt_family="MLT1B"
repeat_region complement(28228. .28522)
/rpt_family="AlusG"
repeat_region complement(28592. .28700)
/rpt_family="MIR"
repeat_region complement(29353. .29537)
/rpt_family="MIR"
repeat_region 29802. .30095
/rpt_family="AlusG"
repeat_region complement(30099. .30219)
/rpt_family="MIR"
repeat_region 30227. .30267
/rpt_family="(TG)n"
repeat_region 30276. .30321
/rpt_family="(CA)n"
repeat_region complement(30322. .30611)
/rpt_family="AluJo"
repeat_region complement(30956. .31151)
/rpt_family="MER3"
repeat_region complement(31533. .31633)
/rpt_family="L2"
repeat_region 31810. .31858
/rpt_family="L2"
repeat_region 32028. .32107
/rpt_family="L1PA4"
repeat_region complement(32919. .33182)
/rpt_family="MIR"
repeat_region 34169. .34314
/rpt_family="CT_rich"
repeat_region 34774. .35007
/rpt_family="MIR"
repeat_region 35589. .35621
/rpt_family="AT_rich"
repeat_region 35763. .36142
/rpt_family="MER41A"
repeat_region 36143. .36452
/rpt_family="AlusG"
repeat_region 36453. .36636
/rpt_family="MER41A"
repeat_region 38011. .38071
/rpt_family="CT_rich"
repeat_region 38083. .38262
/rpt_family="(TTC)n"
repeat_region 38235. .38319
/rpt_family="(TTC)n"
repeat_region 38320. .38499
/rpt_family="(TTC)n"

```

Query Match 2.7%; Score 53.8; DB 9; Length 180385;  
 Best Local Similarity 44.5%; Pred. No. 0.0023;  
 Matches 214; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

```

Qy 305 AGGAGAGAGAAATGCTTGGATCGAGAAATCTTAAAGAACTGACAAATGTCAGAG 364
||||| ||| ||| ||| ||| ||| |||
Db 38536 AGGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38477
Qy 365 CCCAGCTTTCCCAAGAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38476 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38417
Qy 425 GGGATACGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 484
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38416 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38357
Qy 485 CCGAGATGCTGTCTCCACACTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38356 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38297
Qy 545 AGACCAACCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38296 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38237
Qy 605 AGATTGAGCTTCTAATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38236 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38177
Qy 665 GGTGGACACAGTCAGCGGTGGAACAGCTGCTGTGTAATCTGTCTCTCAAGAAAGAGT 724
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38176 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38117
Qy 725 ACAGAAATCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38116 AGCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38057
Qy 785 A 785
Db 38056 A 38056

```

RESULT 38  
 AL589661 241432 bp DNA linear ROD 30-JAN-2002  
 LOCUS Mouse DNA sequence from clone RP23-58B7 on chromosome 15.  
 DEFINITION AL589661  
 ACCESSION AL589661.21 GI:18476659  
 VERSION HTG.  
 KEYWORDS house mouse.  
 SOURCE Mus musculus.  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 REFERENCE 1 (bases 1 to 241432)  
 AUTHORS Dunn, M.  
 JOURNAL Direct Submission  
 Submitted (30-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk  
 On Feb 1, 2002 this sequence version replaced 91:18151481.  
 COMMENT During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30);  
 an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em., EMBL; Sw.,  
 SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP  
 database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep RP23-58B7 is

from the RPCT-23 Mouse PAC Library  
constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBAC3.6  
This sequence is the entire insert of clone RP23-58B7.

FEATURES

source

Location/Qualifiers  
1. 241432

/organism="Mus musculus"

/db\_xref="taxon:10090"

/chromosome="15"

/clone="RP23-58B7"

/clone\_11b="RPCT-23"

misc-feature

147938..148190

/note="Sequence from uni-directional dGTP big dye  
terminator reads only."

BASE COUNT

64293 a 50995 c 55888 g 70256 t

ORIGIN

Query Match

2.7%; Score 53.8; DB 10; Length 241432;

Best Local Similarity 47.4%; Pred. No. 0.0024;

Matches 228; Conservative 0; Mismatches 247; Indels 6; Gaps 2;

QY 305 AGGAGAGGAGAAATGCTTGTGATCGAGATTCTTAAGAATGACTGGACATGTCAGAG 364

Db 147752 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 147811

QY 365 CCCAGCTTCCCGAAAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 424

Db 147812 AACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 147871

QY 425 GGGATACGCTGGAAGAACGCAATGCTACTGTGTATCTTGCAGCAGCGCTTGGGCAAGG 484

Db 147872 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 147931

QY 485 CCGAGATGCTGTGCTCCACACTGAAAAAGCATGAAGTACTTAGACGACGACGAGATG 544

Db 147932 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 147991

QY 545 AGACCAACAAGACCAAGAGGCGGCGCTCAGACGACCAAGTAAAGACCATGTGAGC 604

Db 147992 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 148045

QY 605 AGATTGACTTCTACTCAGACGACGCTCCCTGAGTGGAGAGATGATCCGAGACATGG 664

Db 148046 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 148105

QY 665 GTGTGGACAGTCAAGCGGTGGAACAGCTGCTGTGTACTGTGTCTCTCAAGAAAGAT 724

Db 148106 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 148165

QY 725 ACGGAGATCTTAAAAAGAGGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 784

Db 148166 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 148225

QY 785 A 785

Db 148226 A 148226

RESULT 39

AC051623/c

LOCUS

DEFINITION Mus musculus clone RP23-41205, WORKING DRAFT SEQUENCE, 17 unordered

pieces.

ACCESSION AC051623

VERSION AC051623.4

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 195294)

AUTHORS

McCombie,W.R., Baker,J.P., Bahrel,A., Bal,H., Dedhia,N.N., de la  
Bastide,M., Huang,E.N., King,L., Kirchoff,K.A., Miller,B.,  
Nascimento,L.U., O'Shaughnessy,A.L., Preston,R.R., Rodriguez,M.A.,  
Shah,R.S., Shekher,M., Spiegel,L.A., Tolt,K. and Vill,M.D.

TITLE

Unpublished

REFERENCE

2 (bases 1 to 195294)

AUTHORS

McCombie,W.R.

TITLE

Submitted (15-APR-2000)

Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring  
Harbor, NY 11724, USA

On Sep 6, 2000 this sequence version replaced g1:8493559.

Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor

Laboratory

Center code: CSHL

Web site: <http://www.cshl.org/genseq>

Contact: [mcombie@cshl.org](mailto:mcombie@cshl.org)

Center project name: RP23-41205

Center clone name: RP23-41205

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 17 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 54595: contig of 54595 bp in length

54596 54735: gap of unknown length

54736 79339: contig of 24604 bp in length

79340 79479: gap of unknown length

79480 97456: contig of 17957 bp in length

97457 97577: gap of unknown length

97578 114331: contig of 16755 bp in length

114332 114471: gap of unknown length

114472 127452: contig of 12981 bp in length

127453 127591: gap of unknown length

127592 138079: contig of 10488 bp in length

138080 138218: gap of unknown length

138219 148562: contig of 10344 bp in length

148563 148701: gap of unknown length

148702 157987: contig of 9286 bp in length

157988 158126: gap of unknown length

158127 167394: contig of 9268 bp in length

167395 167533: gap of unknown length

167534 174603: contig of 7070 bp in length

174604 174742: gap of unknown length

174743 178566: contig of 3824 bp in length

178567 178705: gap of unknown length

178706 182160: contig of 3455 bp in length

182161 182299: gap of unknown length

182300 185278: contig of 2979 bp in length

185279 185417: gap of unknown length

185418 188252: contig of 2835 bp in length

188253 188391: gap of unknown length

188392 191030: contig of 2633 bp in length

191031 191169: gap of unknown length

191170 193703: contig of 2534 bp in length

193704 193842: gap of unknown length

193843 195294: contig of 1452 bp in length.

FEATURES

source

Location/Qualifiers  
1. 195294

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="RP23-41205"

BASE COUNT

48944 a 48556 c 46839 g 48701 t 2254 others

ORIGIN





Db 201716 GGAGGAGGAGGAGAGAGAGAAGAAGAAGAGGAGAGAGAGAGCAGAAAGA 201775  
QY 627 CCAGCTCCCTGAGGTGGAGGAGATGATCCGAGCATGGGTGTGGACATGCAGCGGTGA 686  
Db 201776 GAAGGAGGAGGAGGAGGAGAGAGAAGAAGAGAGAGAGAGAGAGAGAGAAAGA 201835  
QY 687 ACAGCTGGCTGTGTACTGTGTCTCTCAAGAAAGAGTACGAGATCTTAAAGAGGCACG 746  
Db 201836 GGAGGAGAGAGAGAGAGAGAGAAAGAAAAGAGAGAGAGAGAGAGAGAGAGGA 201895  
QY 747 GAAGGCCTCAGGGGAGGTGGCTGACAAAGCTGAGGAAGA 785  
Db 201896 GAAGGAGAGAGAGAGAGAGAAAGAAGAAGAAGAAGAAGA 201934

Search completed: September 4, 2002, 14:29:52  
Job time: 13451 sec

**THIS PAGE BLANK (USPTO)**